NAVFAC
Naval Facilities Engineering Command

ENGINEERING SERVICE CENTER
Port Hueneme, California 93043-4370

TECHNICAL REPORT
TR-2318-ENV

DEMONSTRATION AND CERTIFICATION OF AMPHIBIAN ECOLOGICAL RISK ASSESSMENT PROTOCOL FINAL REPORT
ESTCP Project Number ER-0514

by

John Bleiler, ENSR
Dr. David Pillard, ENSR
Dr. Mark Johnson, USACHPPM
Christine Archer, ENSR
Matt Bazar, USACHPPM
Dr. Doris Anders, AFCEE
David Barclift, NAVFAC LANT
Johnny Noles, NAVFAC LANT
Dr. Paula Henry, USGS
Kathleen Harvey, ENSR
Amy Hawkins, NAVFAC ESC

Environmental Security Technology Certification Program

(ESTCP)

Demonstration and Certification of Amphibian Ecological Risk Assessment Protocol
ESTCP Project Number ER-0514

Final Report (revision 2)

Naval Facilities Engineering Service Center
Port Hueneme, California

April 2009
Approved for public release. Distribution is unlimited

Contents

Appendices

List of Tables

List of Figures

Abbreviations and Acronyms

AFB	Air Force Base
AFCEE	Air Force Center for Engineering and the Environment
ANOVA	Analysis of Variance
APG	Aberdeen Proving Ground
ARMI	Amphibian Research and Monitoring Initiative
ASTM	American Society of Testing and Materials
AVS	Acid Volatile Sulfides
CEC	Cation Exchange Capacity
CERCLA	Comprehensive Environmental Response, Compensation, and Liability Act
COPC	Constituent of Potential Concern
COTS	Commercially-Off-The-Shelf
CWA	Clean Water Act
DNT	2,4-Dinitrotoluene
DOC	Dissolved Organic Carbon
DoD	Department of Defense
DOE	Department of Energy
DOI	Department of the Interior
E	Energy
Eco-SSL	Ecological Soil Screening Level
ERA	Ecological Risk Assessment
ESTCP	Environmental Security Technology Certification Program
FCETL	Fort Collins Environmental Toxicology Laboratory
GPS	Global Positioning System
I	Intensity
IACUC	Institutional Animal Care and Use Committee
IDW	Investigation Derived Wastes
IPR	In-Progress Review
IR	Installation Restoration
LCS	Laboratory Control Spike
LOAEC	Low Observed Adverse Effect Concentration
mg/kg	milligram per kilogram

mg/L	milligram per liter
MS/MSD	Matrix Spike/Matrix Spike Duplicate
NAVFAC	Naval Facilities Engineering Command
NAVFAC LANT	Naval Facilities Engineering Command Atlantic
NOAEC	No Observed Adverse Effect Concentration
NAVFAC ESC	Naval Facilities Engineering Service Center
PAH	Polycyclic Aromatic Hydrocarbon
PCB	Polychlorinated Biphenyl
PEC	Probable Effect Concentration
QA/QC	Quality Assurance/Quality Control
QAPP	Quality Assurance Project Plan
RCRA	Resource Conservation and Recovery Act
RDX	Cyclotrimethylenetrinitramine
RPD	Relative Percent Difference
SEM	Simultaneously Extracted Metals
SEM	Standard Error on the Mean
SETAC	Society of Environmental Toxicology and Chemistry
SOP	Standard Operating Procedure
TBD	To Be Determined
TEC	Threshold Effect Concentration
TNT	2,4,6-Trinitrotoluene
TOC	Total Organic Carbon
USACHPPM	United States Army Center for Health Promotion and Preventive Medicine
USEPA	United States Environmental Protection Agency
USFWS	United States Fish and Wildlife Service
USGS	United States Geological Survey
UXO	Unexploded Ordnance
λ	Wavelength
XRF	X-ray Fluorescence

Acknowledgements

This project was supported by the Department of Defense (DoD) through the Environmental Security Technology Certification Program (ESTCP) with funding provided to the Naval Facilities Engineering Service Center (NAVFAC ESC) and its DoD partners Naval Facilities Engineering Command (NAVFAC), Naval Facilities Engineering Command Atlantic (NAVFAC LANT), U.S. Army Center for Health Promotion and Preventive Medicine (USACHPPM) and Air Force Center for Engineering and the Environment (AFCEE), as well as their contractor ENSR Corporation (ENSR).

Executive Summary

This ESTCP project, *Demonstration and Certification of Amphibian Ecological Risk Assessment Protocol, CU-0514*, was designed to demonstrate and validate an innovative technique for the evaluation of potential risks to amphibians in palustrine wetland environments. This technique builds on previous DoD research which resulted in development of a tiered amphibian ecological risk assessment (ERA) protocol as well as laboratory toxicity tests for evaluating potential risks to amphibians due to exposure to contaminated soils and sediments (referred to herein as the soil protocol and sediment protocol, respectively).

The field demonstration described in this technical report was conducted to achieve the following objectives:

- Demonstrate and validate use of the soil and sediment exposure protocols at two DoD sites with potential amphibian risk assessment concerns;
- Apply an amphibian ERA framework at a DoD field site to evaluate whether or not it provides valuable risk management information; and
- Evaluate the use of lead and copper screening values designed to be protective of amphibians developed during the laboratory validation phase of this project (NAVFAC, 2007b).

The soil exposure protocol focused on evaluation of potential risks to terrestrial salamanders, with the red-backed salamander (*Plethodon cinereus*) selected as the model organism. The sediment protocol focused on evaluation of potential risks to early life stage frogs -- the northern leopard frog (*Rana pipiens*) was selected as the test organism for the sediment protocol. Travis Air Force Base (AFB) in California and the Aberdeen Proving Ground (APG) in Maryland were selected as the demonstration sites since both sites have amphibian habitat co-located with contamination associated with firing ranges. Copper and lead concentrations were measured at both demonstration sites; however, copper concentrations in most samples were lower than anticipated and were not present at levels high enough to result in significant adverse impacts to amphibians. Therefore, evaluation of potential impacts due to lead exposure was the primary focus of the field demonstration.

Soil exposure tests were conducted with six samples, including three samples from each of the two demonstration sites. Copper concentrations in all samples were below concentrations associated with effects during the laboratory validation phase of testing so the field demonstration testing focused on evaluation of potential effects on salamanders due to lead exposure. Lead concentrations ranged from 10.8 to approximately 17000 mg/kg in the six tested samples. Survival was not impacted in any of the field demonstration samples, while a concentration of 9,167 mg/kg lead in the laboratory validation phase of testing caused 80% mortality. A reduction in growth was observed over 28 days of exposure to the highest field-collected lead concentration. These results demonstrate that the field-collected soils were substantially less toxic than similar levels of lead in the laboratory spiked soils tested in the laboratory validation phase of testing. Therefore, for the range of conditions encountered at the two demonstration sites, the ecological screening levels derived based on the laboratory validation testing with spiked soils would be sufficiently conservative for use in assessing risks to salamanders exposed to lead under field conditions.

The sediment exposure testing showed similar results with greater toxicity observed in the laboratory validation testing than in the field demonstration testing. A total of 16 sediment samples (eight from each demonstration site) were tested using the 10-day early life stage frog sediment exposure protocol. Lead concentrations ranged from 15 to 17000 mg/kg in the 16 tested samples. Although lead was the primary focus of the demonstration, copper levels in some samples collected from the APG study area may have contributed to observed effects. Impacts on survival and growth of tadpoles were observed in samples collected from both demonstration sites. At both demonstration sites, lead concentrations resulting in adverse impacts were higher than predicted by the sediment screening values developed during laboratory validation testing. These results indicate that, at least for the range of conditions encountered at the two demonstration sites, the screening levels derived using spiked sediments would be sufficiently conservative for use in assessing potential risks to larval amphibians exposed to lead under field conditions.

The tiered amphibian ERA protocol was determined to be useful for conducting both screening level and more sophisticated ERA analyses. The application of the soil and sediment exposure protocols resulted in a more appropriate site-specific assessment of potential risks to amphibians than would have been accomplished using more traditional methods (e.g., comparison to non-amphibian literature-based screening levels, application of alternative soil or sediment toxicity tests using inappropriate receptors like benthic invertebrates or terrestrial worms).

The performance objectives for the field demonstration effort were met. The soil and sediment toxicity testing protocols were appropriate for use at both demonstration sites and were sensitive enough to detect lethal and sub-lethal impacts due to exposure to firing range contaminants. The ERA protocol was also applicable at both demonstration sites. Although results of this testing program have not been submitted to regulatory agencies, the sediment exposure protocol has been applied at several sites under federal and state regulatory review. Technology transfer efforts are on-going and an ASTM standard containing the sediment exposure protocol was approved by ASTM International in November 2007 is currently available on the ASTM website as ASTM E2591-07 Standard Guide for Conducting Whole Sediment Toxicity Tests with Amphibians.

The costs of conducting the tiered amphibian ERA protocol are expected to vary significantly from site to site depending upon the spatial scale of the area under investigation and the number of samples required to meet the data quality objectives. It is anticipated that the cost to conduct the sediment exposure protocol in accordance with the ASTM standard will be within ± 20% of the costs to conduct 10 day sediment toxicity tests with benthic invertebrates. Although the costs to conduct the amphibian sediment toxicity test could be slightly higher than historically used benthic invertebrate sediment toxicity tests, this additional cost is returned when the use of appropriate test species is used to establish remedial goals and avoid unnecessary wetland remediation and restoration.

The sediment exposure protocol and the amphibian ERA framework are both applicable for investigating potential impacts to amphibians at wetland sites under investigation by the DoD or other entities. Although the soil exposure protocol is a valid approach to investigating toxicity from chemicals in soil to a terrestrial salamander, ethical and financial obstacles preclude its regular application as part of site characterization efforts. However, this method may be appropriate for controlled toxicological investigations designed to derive safe soil levels for particular compounds.

1.0 Introduction

The Department of Defense (DoD) Environmental Security Technology Certification Program (ESTCP) has funded the Naval Facilities Engineering Service Center (NAVFAC ESC) and its DoD partners Naval Facilities Engineering Command (NAVFAC), Naval Facilities Engineering Command Atlantic (NAVFAC LANT), U.S. Army Center for Health Promotion and Preventive Medicine (USACHPPM) and Air Force Center for Engineering and the Environment (AFCEE), as well as their contractor ENSR Corporation (ENSR), to demonstrate and validate an innovative technique for the evaluation of potential risks to amphibians in palustrine wetland environments.

This report presents the results of the ESTCP amphibian risk assessment project (*Demonstration and Certification of Amphibian Ecological Risk Assessment Protocol, ER-0514*), with a primary focus on the field demonstration effort. The following reports were previously prepared under the ER-0514 program:

- NAVFAC, 2005. Laboratory Validation Plan for ESTCP Project CU-0514 Demonstration and Certification of Amphibian Ecological Risk Assessment Protocol. December 2005.
- NAVFAC, 2006. Site Selection Memorandum for ESTCP Project CU-0514 Demonstration and Certification of Amphibian Ecological Risk Assessment Protocol. February 2006.
- NAVFAC, 2007a. Field Demonstration Plan for ESTCP Project CU-0514 Demonstration and Certification of Amphibian Ecological Risk Assessment Protocol. May 2007.
- NAVFAC, 2007b. Test Refinement Interim Report for ESTCP Project CU-0514 Demonstration and Certification of Amphibian Ecological Risk Assessment Protocol. June 2007.

1.1 Background

The ESTCP amphibian risk assessment demonstration program is timely and topical, in that nationwide declining amphibian populations and losses of wetland acreage are viewed as significant environmental concerns, and regulatory agencies are increasingly requesting amphibian ERA at DoD sites. In fact, in 2000 the President and Congress directed Department of the Interior (DOI) agencies to undertake a national amphibian research and monitoring initiative (ARMI). This multi-agency effort is coordinated by the U.S. Geological Survey (USGS) and involves numerous state and federal agencies, other organizations, and academic scientists.

The evaluation and remediation of DoD sites requires the careful selection of ERA-based remedial goals. When ERA-based remedial goals are concise, quantitative, and well-supported, the DoD benefits by minimizing costs and expediting response action schedules. Some of the most challenging impediments to regulatory consensus relate to balancing the trade-offs between destructive and costly wetland remediation and leaving residual contamination in place. When selecting remedial goals for addressing wetland contamination, risk professionals are entrusted to balance the objectives of remediation (ecological receptor and habitat protection) with the potential financial costs and short-term and long-term ecological impacts from the disruption caused during remediation. This risk of remedy concern was recently articulated by the United States Environmental Protection Agency (USEPA) in their sediment remediation guidance (USEPA,

2005), which clearly states that relative or net risk analysis is an integral part of the recommended strategy for risk-based closure of sites with sediment contamination concerns.

When selecting appropriate receptors to derive ERA-based remedial goals, amphibians must increasingly be considered. Amphibians play a key ecological role in wetlands, and are an important link in ecological food chains, serving both as predators and prey items. Moreover, public concern regarding recent declines in amphibian populations and additions of many amphibian species to threatened or endangered status suggest that amphibians are important sentinel species in stressed environments. However, because limited ecotoxicity data are available for amphibians, decisions regarding wetland remediation are often inappropriately based on data from aquatic (e.g., fish) or terrestrial (e.g., earthworm) species that are not typical of wetlands, and may be more or less sensitive to chemical stressors than amphibians.

Considerable uncertainty is associated with the application of these terrestrial and aquatic clean-up standards to wetlands. For instance, mineral-based upland soils typically contain low concentrations of organic carbon, whereas wetland hydric soils typically contain much higher organic contents (especially histosols), which have the potential to render many contaminants biologically unavailable as they are bound up in the organic fraction of the soil. Likewise, considerable uncertainty is associated with the application of benthic sediment clean-up standards to wetland systems. Benthic standards are often developed from large lacustrine (i.e., Great Lakes) databases and have little relevance in palustrine wetlands. Although two-phase partitioning (biota and hydric soil) in seasonally flooded wetland soils may influence bioavailability, dynamic three-phase equilibrium partitioning of constituents between water, sediment, and biota, which results in reduced constituent bioavailability in true aquatic sediments, may not always occur in seasonally inundated or saturated palustrine wetland systems.

The DoD is the country's third largest federal land manager, with over 25 million acres at more than 425 military installations. Wetland habitats often comprise a significant portion of open space at DoD facilities, and are prime habitat for various amphibian species, which play a key ecological role in wetlands. This phenomenon is illustrated at the former Naval Air Station in Weymouth, MA where wetlands comprise approximately 40% of the property, and where an amphibian-based ERA was a critical component of the Navy's Installation Restoration (IR) program and ERA-based remedial decisions. Although wetlands comprise a larger percentage of open space at East Coast DoD installations, wetlands containing amphibians are present at installations across the U.S. For instance, at Camp Pendleton, CA, wetland endangered species concerns have played a significant role in the IR program (over 5,700 acres of wetlands and 17 federally listed species occur at this activity), and at least two sites with endangered Arroyo toads (*Bufo microscaphus californicus*) have been remediated.

Wetland remediation can be environmentally destructive, and wetland restoration is publicly sensitive, technically challenging, and financially costly. When inappropriate receptors and methods are used to derive ERA-based remediation goals, the result is often an overestimation of site risks resulting in the unnecessary excavation and destruction of wetlands. In addition to the financial burdens associated with unnecessary wetland remediation, preventable losses of valuable wetland resources may occur if risk-based decisions are based on inappropriate wetland risk assessment techniques. Wetland restoration is still an evolving science, and the short-term and long-term impacts to wetland communities are tangible – wetlands represent successional habitats

that can take decades to mature and many restoration efforts have been unsuccessful. As a result, achieving long-term restoration objectives can be extremely challenging, and decisions relative to wetland alteration should be made with the best available and most current scientific information.

In addition to the direct financial and ecological costs associated with unnecessary wetland remediation and subsequent restoration, many other indirect financial and temporal impacts can arise when inappropriate ERA methods are applied. Potential cost impacts include onsite handling, transport, and disposal of wetland material, which is high in organic content. Saturated organic soils can pose technically challenging problems that have significant cost implications when it comes to treatment and disposal. For instance, excavation of saturated soils can be costly, thermal desorption of wetland soils requires that most of the water be removed, and many hazardous waste landfills will not take a saturated or high organic content soil. Other potential costs include the technical and management efforts associated with reproducing project documents and negotiating with regulatory agencies. One of the more significant overall impacts is the additional time necessary to proceed with remedial design and implementation, and the subsequent delay in site closure and property transfer.

1.2 ER-0514 Work Conducted to Date

This ESTCP project (*Demonstration and Certification of Amphibian Ecological Risk Assessment Protocol, CU-0514*) has been established to build on two previous DoD innovative technology programs:

- NAVFAC recently developed a laboratory toxicity test for evaluating potential risks to early life stage frogs and toads from exposure to sub-aqueous sediments. This technology, referred to herein as the "sediment exposure protocol", evaluates effects on amphibian growth and survival following exposure to contaminated sediments. The final technical report (NAVFAC, 2004) can be downloaded from the Navy's Ecological Risk Assessment homepage (http://web.ead.anl.gov/ecorisk/index.cfm), which is hosted by Argonne National Laboratory. This document also presents the framework for a tiered amphibian ecological risk assessment (ERA) protocol that can be used to assess potential risks to amphibians as part of site evaluations.
- USACHPPM recently completed a similar innovative technology effort resulting in development of a toxicity test focusing on adult salamander exposures to mesic soils. This technology, referred to as the "soil exposure protocol", evaluates effects on salamander growth, survival, and target organ effects based on histopathological evaluations following contaminant exposure (Johnson, et al., 2004).

1.2.1 Protocol Test Refinement

The initial phase of this ESTCP project was focused on validating and refining these existing toxicity test protocols in the laboratory with lead- and copper-spiked sediment and soil prior to the field demonstration. These two constituents were selected for technology validation because they are commonly co-located and are found at military sites and ranges. A primary goal of test validation was to ensure that the amphibian test protocols consider site-specific conditions that

influence exposure (e.g., bioavailability) and yield results that are protective of various life stages of amphibians.

As described in more detail in Section 3.4 (Pre-Demonstration Testing and Analysis), and in previous project deliverables (NAVFAC, 2005; NAVFAC, 2007b), the laboratory validation of the sediment exposure protocol evaluated the impact of a number of bioavailability factors (i.e., pH, total organic carbon (TOC), cation exchange capacity (CEC), grain size) on the observed toxicity of lead and copper. The laboratory validation testing also evaluated the duration of the test by conducting the sediment exposure protocol through frog metamorphosis. The no observed adverse effect concentrations (NOAECs) and low observed adverse effect concentrations (LOAECs) identified during the sediment exposure validation testing can be used as preliminary sediment screening levels.

The soil exposure protocol was evaluated in the laboratory validation phase of the project to assess whether the protocol, originally developed to evaluate energetic compounds, could be used to evaluate inorganic compounds. The laboratory validation testing finalized the test protocol itself relative to the endpoints evaluated and developed dose-response relationships that could be used as soil screening levels (i.e., NOAECs and LOAECs).

The laboratory validation phase of testing finalized the two protocols that were used in the field demonstration effort. The validation testing also developed preliminary sediment and soil screening levels that that could be used to evaluate potential impacts to amphibians due to exposure to copper or lead in the field. The remainder of this document describes the 2006 field demonstration of these two protocols, as well as the preliminary screening levels, at the Travis Air Force Base (AFB) in California and the Aberdeen Proving Ground (APG) in Maryland.

1.3 Objectives of the ESTCP Field Demonstration

The field demonstration effort builds on the technology refinement phase of work, which was described in the December 2005 *Laboratory Validation Plan* (NAVFAC, 2005) and in the June 2007 *Test Refinement Interim Report* (NAVFAC, 2007b). The *Field Demonstration Plan* (NAVFAC, 2007a) described the proposed tasks for the demonstration effort. Appendices A and B from the approved *Field Demonstration Plan* have been re-presented in this document; these appendices respectively present Analytical Methods Supporting the Experimental Design (Appendix A) and the Quality Assurance Project Plan (Appendix B). The site-specific Health and Safety Plan presented in the *Field Demonstration Plan* (NAVFAC, 2007a) is on file at ENSR and available upon request. Appendix C presents a summary of the results from the analytical laboratories and the toxicity testing laboratories.

The specific objective of the Field Demonstration Program is to validate and demonstrate use of the soil and sediment exposure protocols at two existing DoD sites with potential amphibian risk assessment concerns. This validation and demonstration effort was also designed to validate the previously developed amphibian ERA framework and the ecological screening values developed during the laboratory validation phase of this project (NAVFAC, 2007b).

Two demonstration sites were selected where both the soil and sediment exposure protocols could be applied. These two sites have co-located amphibian habitat with lead and copper contamination

associated with firing ranges. Lead, copper, and other metals like antimony and zinc are expected contaminants in firing-range soils due to their presence in bullets and shell casings (Thorbjornsen, and Myers, 2007).

The primary objective of the field demonstration was to evaluate the potential impact of the lead and copper contamination on existing or future amphibian populations. The amphibian ERA field demonstration included conducting laboratory toxicity tests at select stations co-located with chemical analysis sampling stations, performing field surveys to evaluate habitat and amphibian populations, and evaluating site-specific media concentrations relative to screening values developed during the laboratory validation testing (see Table 1-1). Although the May 2007 *Field Demonstration Plan* (NAVFAC, 2007a) proposed field collection of amphibians in order to conduct tissue and histological evaluations, this task did not prove to be feasible due to concerns about the local extirpation of resident amphibian communities at the demonstration sites.

In addition to conducting the field demonstration, a secondary objective of this project was to achieve American Society of Testing and Materials (ASTM) certification for the testing protocols. Following the field demonstration effort it was determined that the focus of the ASTM certification effort would be on the sediment protocol. As discussed in the *Test Refinement Interim Report* (NAVFAC, 2007b), ethical and financial obstacles involved in the soil protocol likely discourage its use for the identification of risks at individual sites. Therefore, the sediment protocol was submitted for review by the ASTM Sediment Assessment and Toxicology sub-committee in January 2007. This draft ASTM protocol was revised in response to comments from sub-committee members and submitted to the full Biological Effects and Environmental Fate committee in August 2007. It was accepted by the committee in November 2007 and was published in December 2007. The guide is presented in Appendix D and is currently available on the ASTM website as ASTM E2591-07 Standard Guide for Conducting Whole Sediment Toxicity Tests with Amphibians.

Following the November 2006 In-Progress Review (IPR) for this project, ESTCP funded an additional set of studies to investigate the differential sensitivities of larval amphibians to copper and lead. This set of studies was designed to assess, compare, and contrast the responses of multiple amphibian species to exposure to two chemical stressors (lead and copper) in hydric soils. The results of the species sensitivity testing are presented and discussed in Appendix E.

1.4 Regulatory Drivers

As a component of site investigation activities, regulatory agencies are increasingly requesting amphibian ERA at DoD sites, as well as at other government and industry-led state and federal environmental sites. Since limited ecotoxicity data are available for amphibians, it can be difficult to effectively evaluate potential impacts to these receptors. Assessments and remedial decisions may be inappropriately based on data from surrogate receptors (e.g., fish or earthworms) that are not typical of wetlands. Application of the soil and sediment exposure protocols at DoD sites should more successfully address requests from regulatory agencies to assess potential impacts to amphibians.

In addition, the use of the soil and sediment exposure protocols to evaluate the potential for adverse impacts to amphibians is consistent with the requirements of the Comprehensive Environmental

Response, Compensation, and Liability Act (CERCLA), the Resource Conservation and Recovery Act (RCRA), and the Clean Water Act (CWA), as well as state-led programs and DoD standards.

1.5 Stakeholder/End-User Issues

ERAs are often conducted in palustrine wetland systems where traditional risk assessment methods (e.g., screening values, toxicity tests) based on non-wetland receptors may not be the most appropriate way to address the potential for risk to amphibians inhabiting the wetland. The field demonstrations conducted for this project provide additional information to stakeholders on the costs, level of effort, and benefits associated with applying the recently developed toxicity testing methods and the amphibian ERA protocol at a site under investigation.

2.0 Technology Description

Two laboratory toxicity testing protocols have been developed during recent DoD projects. Both protocols evaluate potential impacts to amphibians from exposure to sediment or soil. The soil and sediment exposure protocols have recently completed the final stages of laboratory validation and refinement (NAVFAC, 2007b).

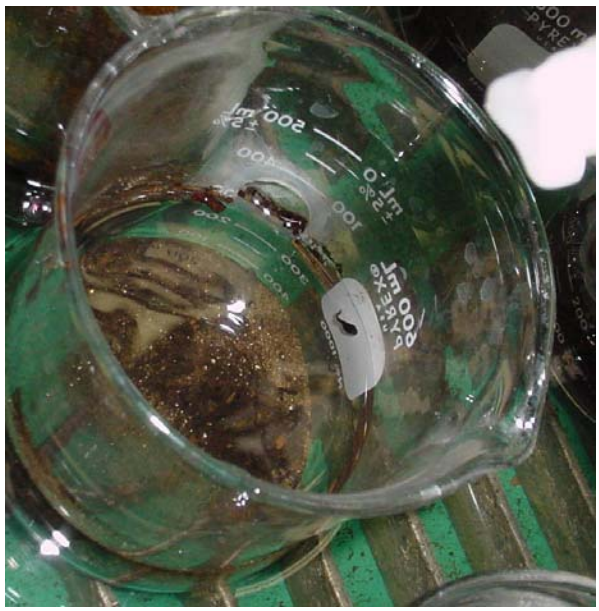
2.1 Technology Development and Application

The sediment and the soil exposure protocols are laboratory bioassays developed to represent model systems for evaluation of amphibian risks on a nationwide basis. Each methodology is summarized briefly below.

The sediment exposure protocol is a 10-day laboratory toxicity test for evaluating potential risks to early life stage frogs and toads from exposure to sediments. This bioassay evaluates effects on amphibian survival and growth following exposure to contaminated sediments. The sediment exposure protocol was developed with a focus on inorganic constituents, and was peer-reviewed and updated to incorporate input from national experts, including DoD, USEPA, USGS, Department of Energy (DOE), and United States Fish and Wildlife Service (USFWS) representatives.

Sediment tests are conducted with recently hatched tadpoles (*Rana pipiens*; Gosner Stages 17-20). Young tadpoles are placed in beakers containing sediment and overlying water (Figure 2-1). The overlying water in each beaker is replaced continuously via a flow-through delivery system. At test termination all living organisms are counted and removed for sub-lethal (width and body length) measurements. Additional endpoints may also be measured at test termination: weight, head-to-vent length, eye width, the occurrence of supernumerary limbs, spinal curvatures, behavioral impairments (e.g., feeding, swimming, orientation), eye displacement. Depending upon project-specific objectives, longer duration studies (i.e., 28 days or until complete metamorphosis) may also be conducted to evaluate potential impacts on tadpole development.

Figure 2-1 *R. pipiens* in Sediment Exposure Protocol Test Chamber



The soil exposure protocol (also presented in Appendix A) assesses adult red-backed salamander (*Plethodon cinereus*) exposure to mesic soils by evaluating effects on salamander growth, survival, and target organs following 28-days of test exposure.

In this protocol, each test organism is placed into an individual Petri dish containing treatment-specific soil (Figure 2-2). Animals are observed at least daily for signs of overt toxicity (e.g., lethargy, sensitivity to touch, abnormal behavior) and body weights are measured weekly. At test termination, surviving salamanders are weighed, anesthetized, and euthanized. Growth, mortality, and health criteria (blood parameters, histological organ evaluation including quantification of liver melanomacrophages) are evaluated as the endpoints for this assay. The liver histopathological biomarkers (melanomacrophages) used in the soil exposure protocol are a non-specific indicator of stress, and show potential as biomarkers for a wide variety of chemical stressors.

Figure 2-2 *P. cinereus* in Soil Exposure Protocol Test Chamber

2.2 Previous Testing of the Technology

Both the soil and sediment exposure protocols are mature technologies, with little remaining development or refinement warranted. Extensive laboratory and data analysis efforts have been conducted during the past four years as part of the research and development of these technologies. Final refinement of both protocols was described in the December 2005 *Laboratory Validation Plan* (NAVFAC, 2005). A final report describing the results of the laboratory validation effort was submitted to ESTCP in June 2007 (*Test Refinement Interim Report*; NAVFAC, 2007b).

The sediment exposure protocol was originally developed under the Navy's YO817 program and presented in the March 2004 *Development of a Standardized Approach for Assessing Potential Risks to Amphibians Exposed to Sediment and Hydric Soils* (NAVFAC, 2004). This document was developed as a guidance manual for risk assessment staff and state/federal regulators involved in the review and approval of risk assessment work plans and reports, and included a standard operating procedure (SOP) for conducting the sediment exposure toxicity test as well as recommendations for field survey methodologies. An initial phase of this ESTCP project included a number of laboratory assays designed to validate the sediment exposure protocol with lead and copper prior to the field demonstration. This laboratory validation phase included the evaluation of several bioavailability factors that could affect the results of the assays. The duration of the test was also assessed. The results of the laboratory validation phase testing were incorporated into the SOP presented in Appendix A.

Since the sediment exposure protocol was developed, it has been used operationally at several state and federal environmental sites, including at the Naval Air Station, Cherry Point, North Carolina, at the Massachusetts Military Reservation, Cape Cod, Massachusetts, at the Naval Weapons Station (NWS) Yorktown, York County, Virginia, at a lead-contaminated state-led site operated by the Massachusetts Highway Department, and at a cadmium-contaminated site led by USEPA Region 4.

The soil exposure protocol methodology was initially established to generate toxicity data for the development of soil screening levels for 2,4,6-trinitrotoluene (TNT), 2,4-dinitrotoluene (DNT), and cyclotrimethylenetrinitramine (RDX). As part of the laboratory validation phase of this ESTCP project, the assay was conducted with copper- and lead-spiked soils to assess how the protocol could be applied to inorganic contaminants. This testing finalized the protocol itself relative to endpoints evaluated for the test metals and developed dose-response relationships that were further evaluated using field-collected mesic soils in the field demonstration.

The refinement testing of both protocols is discussed further in Section 3.4 (Pre-Demonstration Testing and Analysis).

2.3 Factors Affecting Cost and Performance

Implementation of the two toxicity testing protocols according to the established SOPs should be relatively consistent between various laboratories. As with other similar laboratory bioassays, the cost of the assays will be primarily dependent on the number of individual samples tested (i.e., the cost per test may decrease as the number of samples increases) and the level of sub-lethal endpoint assessed. For the sediment assay, measurements on a variety of sub-lethal endpoints (e.g., eye width, behavioral impairments) will be more costly than simply assessing width and length. Extending the duration of the 10 day sediment assay to consider potential impacts on development will also increase costs, but may identify more subtle sub-lethal effects. For the soil assay, extensive histological assessments may increase costs over a more simple evaluation of growth or blood chemistry parameters.

Performance of the assays will be primarily dependent on the supply of test organisms (i.e., laboratory or field-collected) and is seasonally influenced. During the duration of this ESTCP project, the project team continually encountered challenges relative to native test organism

acquisition. Field-collected organisms are generally available from commercial suppliers, but only during the spring months. However, in certain springs (e.g., 2007), factors such as regional droughts adversely affected amphibian populations throughout the southeastern United States, and neither field-collected nor pond-raised amphibian eggs were widely available. Likewise, the availability of laboratory-induced eggs is also seasonally influenced – such eggs are generally available from commercial suppliers from November through April.

Although the SOP and ASTM guide employ standard methods, laboratories with prior amphibian toxicity testing experience should generally be considered for these assays over laboratories that are less familiar with the culturing and testing of amphibians.

2.4 Advantages and Limitations of the Technology

The use of the sediment and soil exposure protocols to assess potential impacts to amphibians in wetlands is often more appropriate than using existing toxicity tests with alternative species. Using aquatic species (e.g., fish), benthic species (e.g., amphipods), or terrestrial species (e.g., earthworms) does not address the unique interaction between amphibians and the sediment or hydric soil milieu. Use of these alternative species may over- or under-estimate potential impacts to amphibians, and may result in less informed risk-management decision-making process.

Due to the potential seasonal availability of amphibians, the use of these protocols may be limited to times of year when the test organisms are available. For the salamander (soil) assay, red-backed salamanders are generally available for testing in the late winter and spring months (February through May). Frog eggs are generally available from commercial vendors during the spring months (field-collected), as well as during the late fall and winter months (reproduction artificially induced in the laboratory). During the refinement stage of this ESTCP program, the project team experienced significant shipment-related mortality during the winter months (possibly due to frog eggs being exposed to extreme winter weather conditions during shipment). During the field demonstration effort and the supplemental species sensitivity testing, delays were incurred due to the availability of frog eggs, as well as other larval amphibians under investigation.

Lastly, the salamander protocol uses significant numbers of field-collected adult organisms. Although the Maryland populations used in the current ESTCP program are robust and do not appear to be substantially affected by the field collection activities in support of this ESTCP program, field collection of adult organisms should only be conducted if local amphibian population and meta-population dynamics are robust enough to support the loss of several dozen adult salamanders.

Proposing and thus promoting the use of this assay to investigate toxicity of mixtures at individual sites risks local and possibly wide-scale extirpation of the species. Additionally, there is circumstantial evidence that these species are relatively long-lived (~ 20 years), adding to the ethical concerns from harvesting these species for site-specific toxicological investigations. Moreover, the test methods used are quite expensive, and likely not feasible for site-specific analysis. Altogether, current constraints suggest that these methods may be appropriate for controlled toxicological investigations designed to derive safe soil levels, but are not feasible for the wide-scale use in determining toxicity from mixtures at individual sites in support of environmental restoration.

These limitations have excluded the soil exposure protocol from being considered as a testing procedure in the development of an ASTM guide to conduct whole sediment toxicity tests with amphibians.

3.0 Demonstration Design

3.1 Performance Objectives

Performance objectives are a critical component of the overall demonstration plan since they provide a measurable basis for evaluating the performance and costs of the technology. Meeting these performance objectives is essential for successful demonstration and validation of the technology.

In general, the quantitative performance objectives for typical remediation-related ESTCP projects (e.g., end-point criteria, remediation time, and analytical sensitivity) are indirectly associated with the performance objectives of this project (e.g., ecological risk and toxicity based performance objectives). Table 3-1 presents the performance objectives for evaluating the field demonstration effort. The evaluation of these objectives will be discussed in Section 4.

Table 3–1 Performance Objectives

Type of Performance Objective	Primary Performance Criteria	Expected Performance Metric
Qualitative	Sediment protocol is applicable to evaluating copper and lead in palustrine wetlands	Correlation between sediment concentrations and lethal or sub-lethal results
	Soil protocol is applicable to evaluating copper and lead in forested uplands	Correlation between mesic soil concentrations and lethal or sub-lethal results
	Collection and biological evaluation of native salamanders is applicable for evaluating potential impacts due to metals	Correlation between mesic soil concentrations and histopathological evaluation
	Regulatory acceptance of toxicity test protocols	Results are accepted by agency as component of ERA
	Versatility of the overall ERA protocol	Application of the ERA protocol at both field demonstration sites
	Technology transferred to other potential end-users	Presentation at conference or in journal; ASTM certification

Table 3–1 Performance Objectives (continued)

Type of Performance Objective	Primary Performance Criteria	Expected Performance Metric
Quantitative – Sediment Exposure Protocol	Sediment toxicity test is valid and acceptable	Mean survival in laboratory control is >80%
	Lethal endpoint indicates toxicity or lack of toxicity	Statistical difference between survival in control or reference samples and site samples
	Sub-lethal endpoints indicate toxicity or lack of toxicity	Statistical difference between sub-lethal endpoints in control or reference samples and site samples (may include growth, abnormalities, behavior, metamorphic stage, or other measurements)
Quantitative – Soil Exposure Protocol	Soil toxicity test is valid and acceptable	Mean survival in laboratory control is >80%
	Lethal endpoint indicates toxicity or lack of toxicity	Statistical difference between survival in control and site samples
	Growth endpoints indicate toxicity or lack of toxicity	Statistical difference between growth endpoints in control or reference samples and site samples
	Blood parameters indicate toxicity or lack of toxicity	Statistical difference between blood parameters measured in control or reference samples and site samples

3.2 Selecting Test Sites

Two potential demonstration sites were selected primarily based on the known presence of amphibian habitats overlapping with copper and lead contamination. Copper and lead were selected as the constituents for technology refinement because they are commonly co-located and are often found at military sites and ranges.

Eleven potential demonstration sites were initially considered in the site selection process. The list of demonstration sites was finally narrowed to two locations based on the likely presence of amphibian habitat with copper and lead contamination, as well as conversations with site personnel on the feasibility of performing the demonstration tasks at the particular locations. The selection criteria used to identify the two sites are briefly described below, and were presented in more detail in the February 2006 *Site Selection Memorandum* (NAVFAC, 2006).

3.2.1 Chemical Parameters

Based on the testing data and screening values generated from the test refinement phase of work (NAVFAC, 2007b), the preferred range of copper levels in mesic soil and palustrine hydric soil was determined to be from 150 milligrams per kilogram (mg/kg) to 6000 mg/kg. The preferred

range of lead levels in mesic soil and palustrine hydric soil is from 100 mg/kg to 3000 mg/kg. These values represent lead and copper concentrations that are often found at DoD sites with firing ranges, and they bracket the upper and lower limits of amphibian mortality and sub-lethal effects, based on the environmental toxicity data generated through this and other amphibian assessment projects. These levels were selected to be high enough to result in observable impacts to the test subjects, but not so high as to result in complete mortality.

The site selection process also considered the presence of other chemical stressors (e.g, pesticides used in grounds maintenance or polycyclic aromatic hydrocarbons (PAHs) from parking areas). Sites with mixed contamination that could make interpretation of the results difficult were eliminated from consideration. Chemical analysis of sediment/soil selected for toxicity testing included other potential stressors in at least a sub-set of the tested samples to assess the potential for interferences.

3.2.2 Ecological Parameters

The presence of suitable habitat for either Plethodontid salamanders or Anurans (frogs or toads) was critical to the site selection process since the protocol focused on evaluating wetland habitats for potential impacts to amphibians and assessing the need for potential wetland remediation to improve existing habitats. However, it was not critical that the selected sites have a reported presence of species from either group, since it is possible an amphibian survey has not been conducted, that the native population may contain related con-generic or con-specific representatives, or that current site conditions have impacted historic populations.

The site selection process eliminated sites with non-chemical stressors in the vicinity of the study area that could make interpretation of the results difficult. Non-chemical stressors included roadways, bridges, drainage ditches, or other physical stressors that might impact the wetlands or the amphibian populations.

Preference was given to sites that did not have known occurrences of federal or state-listed threatened or endangered species in order to reduce the need for federal or state resource agency coordination prior to on-site amphibian collection. However, the presence of a threatened or endangered species did not eliminate a site from consideration.

3.2.3 Site Historical and Logistical Parameters

For the field demonstration to be most applicable and of use to site managers, preference was given to sites currently under investigation to determine whether remedial response actions are required to address potential risks to amphibians. However, otherwise acceptable sites were not eliminated from the process if amphibian impacts had not previously been investigated.

However, some level of previous investigation was important in order to identify levels of lead and copper which may overlap with acceptable amphibian habitat and to identify other stressors (chemical or non-chemical) that might confound the interpretation of the test results.

Safety, accessibility, and the geographic location of the site were also considerations in the site selection process. Safe access to the study area (i.e., wetland habitat) within the DoD facility was

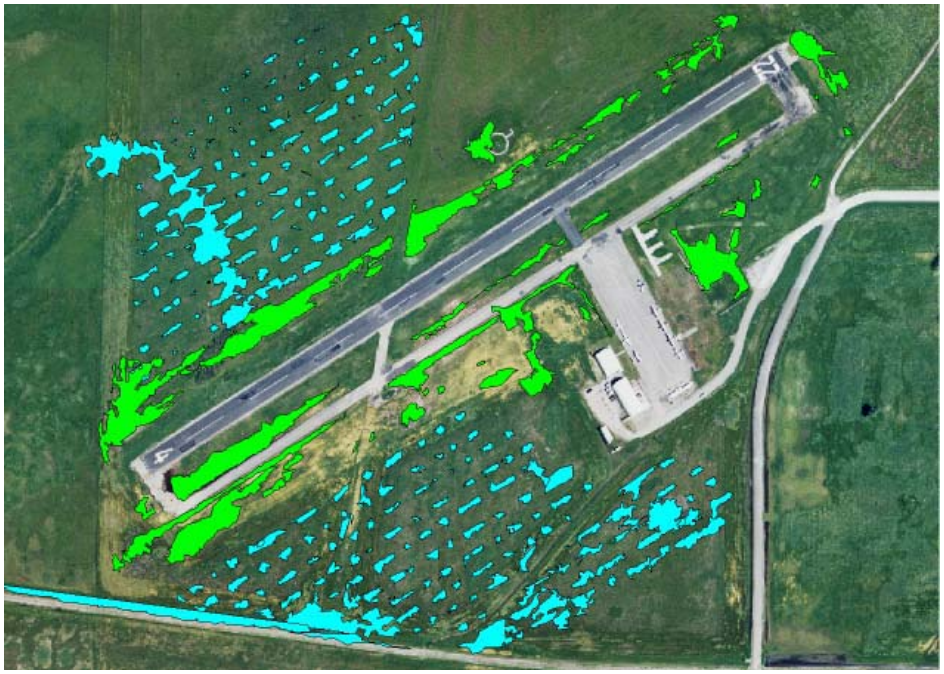
critical. Possible hazards included the potential presence of unexploded ordnance (UXO), active firing ranges, or exposure to on-site hazardous wastes, as well as more general health and safety concerns (e.g., trip hazards, ticks). It was also preferred that the selected sites be representative of the regional diversity of DoD sites, and not both be located in the same geographic region of the country.

3.3 Test Site Description

Two sites were selected for the implementation of the field demonstration phase of work for this project. A brief synopsis of information about these sites and the rationale behind the selection are discussed below.

Site 1: Travis Air Force Base (AFB), Fairfield, California is located midway between Sacramento and San Francisco in northern California. Travis AFB contains several well documented palustrine wetland complexes and vernal pools in close proximity to firing ranges. In addition, Travis AFB has a documented vernal pool complex in close proximity to an active skeet range. Previous vernal pool work at the site has focused on preserving and restoring the natural vernal pool ecosystem, in part to prevent the extinction of a federally endangered plant, the Contra Costa goldfields (*Lasthenia conjugens*). Over 250 vernal pools have been constructed at the AFB to supplement the existing natural vernal pools (see Figure 3-1).

Figure 3-1 Constructed and Natural Vernal Pools at Travis Air Force Base Aero Club Site[1]



[1] Blue represents constructed vernal pools and green represents natural vernal pools. Figure downloaded from website of Sharon Collinge at the University of Colorado – Boulder http://spot.colorado.edu/~sharonc/vernalpools.html

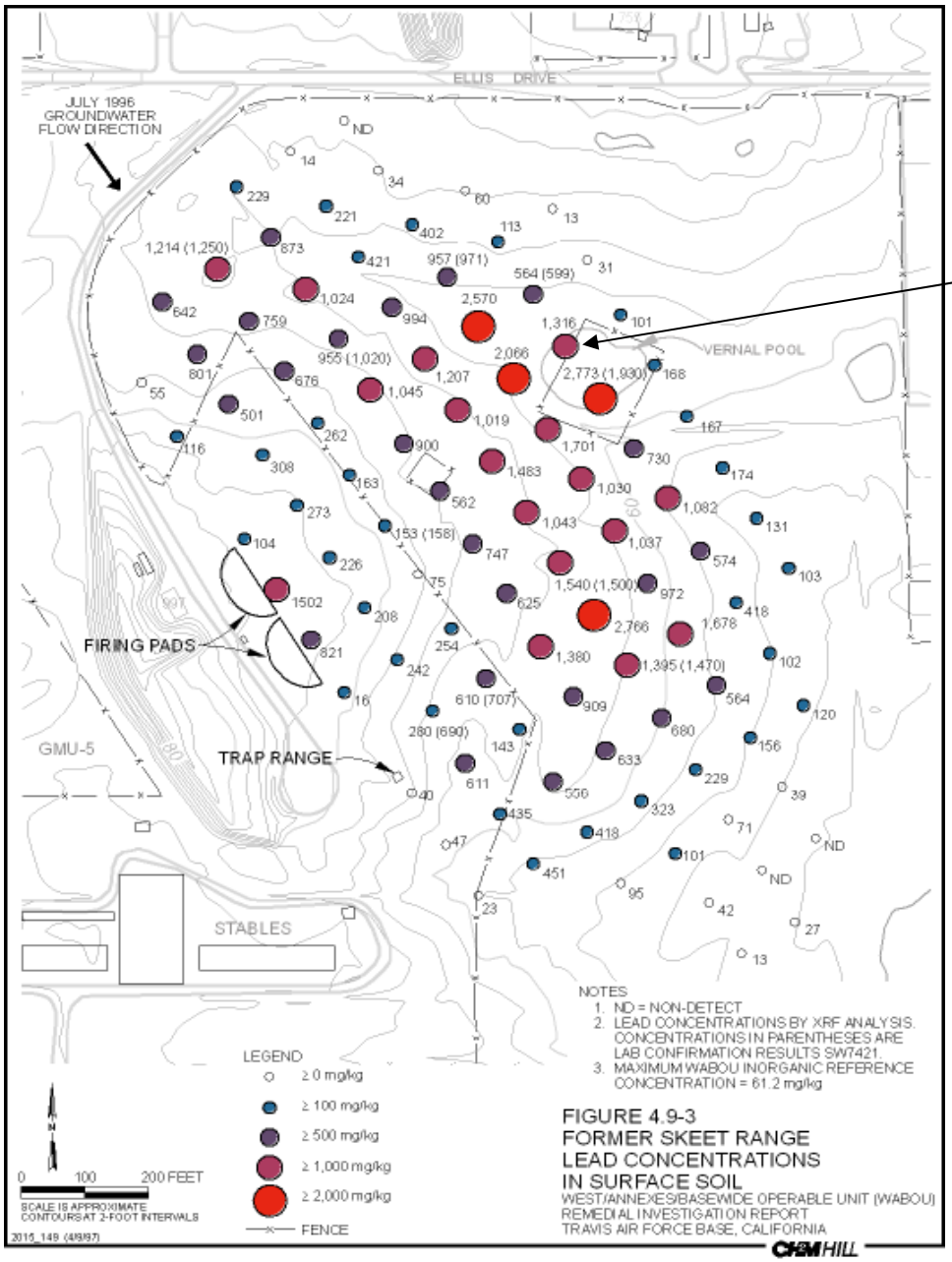
Travis AFB was selected as a preferred location for the demonstration of the amphibian risk assessment protocol based on the variety of vernal pools (both constructed and natural) and palustrine wetlands, the level of interest from the environmental manager, and the likelihood of contamination due to firing ranges.

A brief site reconnaissance visit at Travis AFB was conducted in February 2006 by Navy and contractor personnel with oversight from Air Force personnel. Sites surveyed included the decommissioned small arms firing range and the skeet shooting range, both known to contain elevated levels of lead. Each site was surveyed to determine the presence or absence of potential amphibian breeding habitat. Suitable habitat for amphibians was not present within the area of the firing range that was surveyed for lead and copper soil concentrations. Many small vernal pools exist within the site but none hold enough water for a long enough period of time to be considered suitable amphibian breeding habitat.

A vernal pool located approximately 800 feet northeast of the skeet shooting range was surveyed during the site reconnaissance visit. This vernal pool is down range, down wind, and down slope of the skeet shooting range and is know to contain lead-contaminated soils (pers. com. Glenn R. Anderson, Base Hydrologist, Travis AFB, February 8, 2006) (Figure 3-2). The site is also close enough to the skeet shooting range that lead shot could potentially reach the site. The vernal pool is located within the Travis AFB Equestrian Center horse pasture and the deepest portions of the vernal pool are fenced in to keep the livestock out. The vernal pool, which was inundated at the time of the survey, provides potential breeding habitat for amphibians. Average pool depth was approximately 50 cm with deeper areas over 80 cm. The surface area of the pool was approximately 1,200-square meters. Plant species within the pool were similar to those in the vernal pool southwest of the firing range. A Pacific tree frog (*Hyla regilla)* egg mass and several Pacific tree frog tadpoles were observed within the pool, and the calls of at least one adult tree frog were also heard.

The vernal pool associated with the skeet shooting range was selected as the primary study area at Travis AFB.

Figure 3-2 Vernal Pool Study Area at Travis Air Force Base



Vernal pool study area

Site 2: Aberdeen Proving Ground (APG), Maryland was also selected as a location for the field demonstration of the amphibian risk assessment protocol. One of the ESTCP team partners (Dr. Mark Johnson) is stationed at the proving ground and is intimately familiar with the overlap between amphibian habitat and lead contaminated ranges at this facility.

The facility occupies more than 72,500 acres in Harford County, Maryland and is bounded by the Susquehanna and Gunpowder Rivers, the Chesapeake Bay, and the Amtrak Railroad. APG comprises two principal areas, separated by the Bush River: the northern area known as the Aberdeen Area; and the southern area, formerly the Edgewood Arsenal, known as the Edgewood Area. Activities at the APG have included environmental and chemical research, as well as testing of field artillery, weapons, and ammunition. Numerous exterior and interior firing ranges, automotive courses, and underwater explosive test ponds are located on-site. Due to the active and classified nature of the APG, aerial images of the proposed study area are not included in this report.

The APG also provides large areas of natural habitat for many species. Excluding wetlands within the open water areas, the wetlands at APG total about 13,600 acres or about 35 percent of the land surface area. Non-tidal wetlands total over 6,000 acres with approximately 1,770 acres of emergent wetlands, 4,350 acres of forested wetlands and 134 acres of scrub/shrub wetlands.

More than 40 species of reptiles and amphibians occur within the streams, ponds, wetlands, and forests of the APG. The most abundant amphibian species include bullfrog (*R. catesbeiana*), green frog (*R. clamitans*), northern cricket frog (*Acris crepitans*), northern spring peeper *Pseudacris crucifer crucifer*), southern leopard frog (*R. sphenocephala*), Fowlers toad (*B. fowleri*), and red-backed salamander (*P. cinereus*).[2]

Based primarily on the wide range of wetlands present on-site, the observations of amphibian populations, and the likelihood of contamination due to firing ranges, the APG was selected as a preferred location for the demonstration of the amphibian risk assessment protocol. Army personnel identified an on-site small arms range adjacent to a palustrine wetland as the specific area of study. Although data and mapping are available for the lead contamination in the palustrine wetland complex at the small arms range and the data can be used to select sampling locations, these figures are not currently cleared for public distribution.

3.4 Pre-Demonstration Testing and Analysis

As described in Section 2.2, laboratory validation of the soil and sediment exposure protocol technologies was conducted using soils and sediments spiked with known concentrations of copper and lead. The validation testing conducted for the sediment exposure protocol also evaluated different test durations and several bioavailability factors that could potentially affect the results of the assays. The results of the soil and sediment testing were presented in the *Test Refinement Interim Report* (NAVFAC, 2007b) and were used to finalize the SOPs for each test protocol. The

[2] Information obtained from Aberdeen Proving Ground Ecology website http://www.apg.army.mil/apghome/sites/directorates/restor/ecology.html

results of the field demonstration testing described herein have been compared against the laboratory validation results to determine whether site-specific factors (e.g., organic carbon) or the aging of the metals in the environment has an effect on the toxicity of lead to amphibians.

3.4.1 Soil Exposure Protocol

Table 3-2 presents a summary of the survival and growth of salamanders exposed to copper- or lead-spiked soils during the laboratory validation phase of testing. The 28-day copper exposure assay indicated no survival effects at 803 mg/kg; however, reduced survival was observed at 1,333 mg/kg and above. Eleven of twenty salamanders in the 1,333 mg/kg group were found dead or were euthanized prior to the scheduled test termination due to humane considerations; death or euthanasia occurred from 1 to 3 days post-exposure. All twenty of the salamanders in the 2,700 mg/kg group were found dead within 24 hours of initial exposure. All other salamanders survived to test termination at Day 28. There were no significant differences among treatments compared to controls for body weight, or for white and red blood cell counts (leukocytes and erythrocytes, respectively). There were no copper-related histopathologic findings.

Table 3–2 Salamander Survival and Growth Following 28 Days of Exposure to Copper- or Lead-Spiked Soil

	Concentration (mg/kg) [1]	Mean Survival (%)	Mean Body Weight (g)
Copper	18	100	0.7847
	283	100	0.8050
	803	100	0.8287
	1333	45	0.7174
	2700 [2]	0	NA
Lead	14	100	0.8977
	553	100	0.8988
	1700	100	0.9151
	4700	85	0.9026
	9167	20	0.7668

Survival LOAEC is indicated in boldface text
1 – Concentrations based on the average measured value from three sampling events (Day 0, Day 14, and Day 28).
2 – Concentration is based on one sampling event due to complete mortality prior to second sampling event.

The 28 day lead exposure assay indicated no survival effects at 1,700 mg/kg, but reduced survival in the 4,700 mg/kg and 9,167 mg/kg treatments. Three of twenty salamanders in the 4,700 mg/kg group and sixteen of the twenty salamanders in the 9,167 mg/kg group were found dead or were euthanized prior to the scheduled test termination due to humane considerations; death or euthanasia occurred from 1 to 5 days post-exposure. All other salamanders survived to test termination at Day 28. Although the mean body weight of the salamanders exposed to the 9,167

mg/kg treatment is lower than in the control treatment; this difference was not reported to be statistically different. No significant differences among treatments compared to controls were observed for erythrocyte counts and hemoglobin levels. However, the 4,700 mg/kg treatment had a low leukocyte count compared to controls (p=0.007). There were no lead-related histopathologic findings.

Although only a relatively small dataset was available for the salamander assay, the data provide an important preliminary estimate of potentially toxic levels of copper and lead. The no observed adverse effect concentrations (NOAECs) and low observed adverse effect concentrations (LOAECs) for lethal and sublethal endpoints were used to generate preliminary soil screening levels that could be used to evaluate potential impacts to amphibians due to exposure to copper or lead in the field (Table 3-3).

Table 3–3 Summary of Soil Screening Values Developed During Laboratory Validation Testing

Compound (mg/kg)	Survival		Sub-Lethal Endpoint	
	NOAEC	LOAEC	NOAEC	LOAEC
Copper	803	1333	2700 [1]	>2700 [1]
Lead	1700	4700	1700 [2]	4700 [2]
1 - No sub-lethal effects were observed at the highest tested concentration.				
2 - Reduced leukocyte count observed at 4700 mg/kg lead.				

The laboratory validation testing conducted for the soil exposure protocol indicated that adverse lethal or sub-lethal effects would not be expected at approximately 800 mg/kg copper or 1,700 mg/kg lead (the survival NOAEC). Sites with soil concentrations at or below these levels would not warrant additional evaluation for impacts to amphibians. Soil concentrations of 1,333 mg/kg copper and 4,700 mg/kg lead (the survival LOAEC) may result in reduced amphibians survival and sites with concentrations at and above these levels likely warrant a more detailed assessment of metals bioavailability and toxicity. The sub-lethal NOAECs and LOAECs were also used to derive more conservative screening levels; however, few sub-lethal impacts were observed in the testing so these levels may be overly conservative.

These preliminary screening values are likely to be overly protective since the metals used in the laboratory testing were selected to represent a maximum exposure and may be more bioavailable than metals found within more natural soils where organic carbon and the aging process may limit bioavailability. In addition, it is difficult to know the significance of the impact that a reduced leukocyte count would have on the health of an individual salamander or a salamander population exposed to lead in the soil. Basing a sub-lethal screening level on this endpoint may be overly protective of population-level effects.

3.4.2 Sediment Exposure Protocol

The laboratory validation testing for the sediment exposure protocol was conducted through a sufficient range of copper and lead concentrations and test durations to result in a range of lethal

and sub-lethal responses. The evaluation of bioavailability variables such as TOC, pH, grain size and CEC, provided valuable information on the important impact these chemical and physical factors may have on modifying the bioavailability, and therefore toxicity, of a contaminant in the field.

The NOAECs and LOAECs derived during the laboratory validation testing may be appropriate preliminary screening values for evaluating potential risks to amphibians due to exposure to copper or lead. Table 3-4 presents preliminary screening values based on the NOAECs and LOAECs derived from 10 day tests with copper- or lead-spiked sediments. The bioavailability variables were not modified in these tests so these screening values would be most applicable to sediments that are similar to the base sediment used by the laboratory (sediment with low TOC [0.066%] and pH of approximately 7.5).

Table 3–4 Summary of Sediment Screening Values Developed During Laboratory Validation Testing

Compound (mg/kg)	Survival		Sub-Lethal Endpoint [1]	
	NOAEC	LOAEC	NOAEC	LOAEC
Copper	230	>230	21	87
Lead	1200	>1200	100	260
Values were derived from 10-day tests conducted using laboratory control sediment spiked with either copper or lead. No bioavailability factors were modified. The TOC level in the base sediment was 0.066%. Concentrations based on measured values at test initiation. 1 – Based on lower of values for body width and body length endpoints.				

Sediment concentrations above the laboratory–derived LOAECs may require additional investigation and may present a potential ecological risk to amphibian receptors. Concentrations below the NOAECs are unlikely to cause harm to the local amphibian population and these sites may require no further investigation.

The validation testing indicated that the copper and lead dilution series studies, as well as several of the studies where TOC and pH were modified, were most consistently related to the observed effects. No apparent relationships were observed between toxicity and copper or lead concentrations when grain size and CEC were included in the regression models. Table 3-5 presents the survival data extracted from several studies with varying levels of TOC in sediments spiked with either copper or lead (mixture data are not considered in this table). The general trend is that when metals concentrations are elevated, more toxicity (i.e., lower survival) is observed in sediments with lower TOC levels, particularly below 1%. For example, in the copper tests a concentration of approximately 250 mg/kg results in 60% survival in sediment with <1% TOC but survival increases to 100% at approximately 4% TOC. A similar pattern occurs in the lead tests where 780 mg/kg lead results in only 65% survival in sediment with <1% TOC; however, when the TOC is increased to 10%, concentrations as high as 1367 mg/kg do not adversely impact survival.

Table 3–5 Influence of Total Organic Carbon Content on Tadpole Survival During Laboratory Validation Testing

% TOC Range	% TOC	Copper-Spiked Sediment Tests		Lead-Spiked Sediment Tests	
		Copper [1] (mg/kg)	Mean Survival (%)	Lead [1] (mg/kg)	Mean Survival (%)
<1% TOC	0.06	7.2	100	4.5	97.5
	0.06	7.2	100	4.5	97.5
	0.06	15	100	63	100
	0.06	21	97.5	100	100
	0.06	87	97.5	260	100
	0.06	140	95	680	100
	0.06	230	77.5	1200	70
	0.64	39 [2]	0	780	65
	0.64	250	60	1500	35
1 - 4% TOC	1.4	80	100	800	95
	1.4	140	90	1200	80
	3.9	143	100	933	100
	3.9	250	100	1400	95
>10% TOC	10.2	142	85	911	100
	10.2	248	90	1367	95

1 - Concentrations based on measured values at test initiation.

2 - Copper concentration in some replicates of this treatment may have been higher than the reported value of 39 mg/kg from a sample collected prior to the start of the test. The target concentration was 350 mg/kg and copper levels in other sediments using the same spiked sediment ranged from 140 to 248 mg/kg (average was 169 mg/kg). Total mortality at this copper concentration is unusual.

It is clear that the same level of a contaminant may be more or less toxic depending on the environmental conditions (e.g., TOC) of the sampling location.

3.5 Testing and Evaluation Plan

The following sub-sections describe the details of the on-site activities conducted during the field demonstrations at Travis AFB and the small arms range at APG. Since lead was the primary contaminant present at these firing ranges, the focus of the field sampling was to identify and collect samples that would represent a range of lead concentrations. Copper was also measured, but historic information and initial reconnaissance samplings indicated that copper levels were not likely to be present at levels high enough to result in significant adverse impacts to amphibians. Therefore, potential impacts due to lead exposure became the focus of the field demonstration.

3.5.1 Demonstration Set-Up and Start-Up

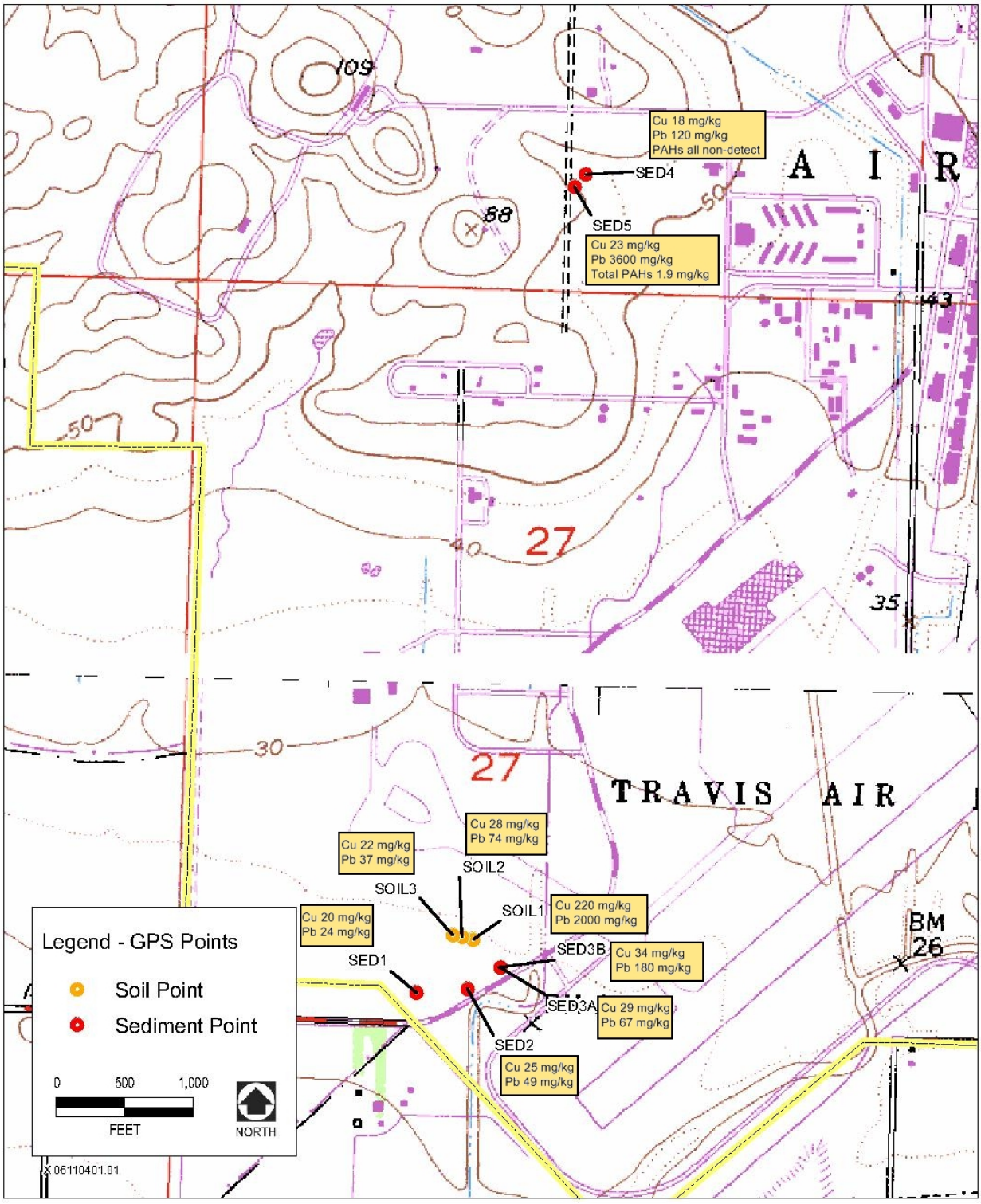
Field reconnaissance visits were conducted at both of the selected sites prior to the field demonstration. Initial field reconnaissance site visits by members of the project team were conducted in February 2006 at Travis AFB and in April 2006 at APG. These field reconnaissance visits helped to determine approximate sampling locations, access, and other logistics.

Field instruments used at each site included a global positioning system (GPS) unit, echosounder, and YSI multi-parameter meter. Calibration procedures for laboratory instruments consisted of initial calibrations, initial calibration verifications, and continuing calibration verifications.

Prior to collecting soil or sediment samples, an X-ray fluorescence (XRF) survey was conducted by USACHPPM personnel at each demonstration site to identify appropriate sampling locations. XRF is a non-destructive qualitative and quantitative analytical technique used to determine the chemical composition of a sample. An element is identified by its characteristic X-ray emission wavelength (λ) or energy (E). The amount of an element present is quantified by measuring the intensity (I) of its characteristic emission. The XRF survey information was used to select soil/sediment sampling locations such that they represented a concentration gradient bracketing and containing the concentrations suspected to result in lethal and sub-lethal responses in amphibians, based on the previously conducted work (including the laboratory refinement phase of this ESTCP project).

Because few lead samples were previously analyzed in the vicinity of the Travis AFB vernal pools, the project team collected nine surficial soil and sediment samples during the February 8, 2006 site reconnaissance (Figure 3-3). All sampling locations were marked in the field with surveyors flagging and GPS data were collected. Samples were packaged and shipped to Paragon Analytics of Fort Collins, Colorado for copper, lead, and PAH analyses (PAHs were evaluated to ensure that there was no residual organic contamination associated with an adjacent railroad track). The analytical data from the field reconnaissance effort are presented in Figure 3-3.

Figure 3-3 Travis Air Force Base Site Reconnaissance Sampling Locations and Analytical Data



The Travis AFB reconnaissance also included an ecological survey of the decommissioned small arms firing range and the skeet shooting range to determine the presence or absence of potential amphibian breeding habitat. The surveys indicated that, assuming that the vernal pools observed on site persist through the breeding season and long enough for the larvae to metamorphose, suitable amphibian breeding habitat exists in the vicinity of both the small arms firing range and skeet shooting range sites. However, no suitable breeding habitat was identified within the area of the small arms firing range that had been surveyed for copper or lead contamination. Based on the results of the analytical chemistry sampling effort and the habitat survey, the Travis AFB demonstration effort focused on the vernal pool in the vicinity of the skeet shooting range. An XRF survey was conducted by USACHPPM personnel one day before the field demonstration sampling in order to focus the sediment and soil sampling in the vicinity of the vernal pool itself. Due to the active nature of the skeet shooting range, communication with skeet range personnel was critical. No other safety issues were noted during the field reconnaissance effort.

The lead contamination in the APG palustrine wetland at the small arms range is well documented; therefore, the project team determined that no additional field sampling was warranted as part of the reconnaissance. On April 10, 2006, just prior to the field demonstration effort, an XRF survey was conducted by USACHPPM personnel in order to identify soil and sediment sampling locations. The area was also cleared with a magnetometer to allow for safe digging. The XRF survey identified a range of lead concentrations in soil and sediment from 33 mg/kg in a reference location to 12,387 mg/kg in a soil sample. These data were used to select sampling locations for the field demonstration effort. The sampling area was limited to the area cleared of UXO by Army personnel. No other safety issues were noted during the field reconnaissance effort.

3.5.2 Period of Operation

A brief site reconnaissance was conducted at Travis AFB in February 2006 by Navy and contractor personnel with oversight from Air Force personnel. The field sampling effort was conducted on March 27 and 28, 2006. An XRF survey was conducted on March 27, 2006 in order to focus the sediment and soil sampling planned for the next day. Surface water samples were collected from the vernal pool and the reference location on March 27, 2006 and were submitted for chemical analyses. Sampling on March 28, 2006 involved the collection of soil and sediment samples for chemical analyses and toxicity testing.

An XRF survey of the APG study area was conducted by USACHPPM personnel on April 10, 2006 in order to identify soil and sediment sampling locations. The APG field sampling effort was conducted on April 12, 2006 and involved the collection of soil and sediment samples for chemical analyses and toxicity testing and surface water samples for chemical analyses.

Due to a lack of available frog eggs, there was a delay in the initiation of the sediment toxicity testing program. The sediment testing was conducted by ENSR's Fort Collins, Colorado Environmental Toxicology Laboratory (FCETL) from December 18 to 28, 2006 for the APG samples and from January 13 to 23, 2007 for the Travis AFB samples. It is recognized that it is desirable to initiate sediment tests as soon as possible following field collection of sediments. However, due to the lack of eggs this was not possible. It was also not possible to conduct additional field efforts to collect fresh sediments once eggs became available. All samples were stored in the dark at 4°C to minimize any changes to the sediment. One concern with longer

holding times is the loss of some labile chemicals such as ammonia and volatile organic compounds that can degrade or volatize during storage. However, more stable sediments can be stored for much longer periods of time with little change in toxicity (USEPA, 2001). Since the compounds of interest at these two sites were metals, the concentrations were not expected to change significantly over time.

This was confirmed by re-analyzing the copper and lead levels in the sediments prior to the sediment toxicity tests. Copper concentrations at the start of the test ranged from 68% to 124% of the concentrations measured just after sampling with the largest variation noted in samples with low copper concentrations (i.e., 17 mg/kg after field collection to 21 mg/kg at test initiation). Less variability was observed for the lead concentrations with concentrations at the start of the test ranging from 76% to 113% of the concentrations measured just after sampling. For both metals, the change in concentrations over time was not consistently higher or lower than that observed initially. The sediments at the time of test initiation were determined to still be appropriate to use for testing; however, it is possible that unmeasured changes occurred during the extended holding time that may have had an impact on the outcome of the tests.

The soil exposure testing was conducted by USACHPPM's Aberdeen, Maryland laboratory between May 24 and June 23, 2006 with tests starts staggered such that each test would run for 28 days.

A protocol deviation in the soil testing resulted in half of the salamanders being exposed for 29 days instead of 28 days. The necropsies were originally planned over the course of two days with the beginning exposures staggered accordingly; all were necropsied on one day. Therefore, half the animals were exposed for an extra day.

Sediment, soil, and surface water analyses (e.g., TOC, grain size, metals) and analytical data review were conducted between March and October 2006.

3.5.3 Amount/Treatment Rate of Material to be Treated

This section is not applicable to this field demonstration project.

3.5.4 Residuals Handling

Following the completion of sampling, residual soils and sediments were discarded adjacent to the sampling location. Any investigation derived wastes (IDW) generated during the decontamination of equipment were containerized and disposed of according to the DoD facility's instructions.

3.5.5 Operating Parameters for the Technology

This section is not applicable to this field demonstration project.

3.5.6 Experimental Design

The primary objective of the field demonstration was to evaluate whether or not the sediment and soil exposure protocols and the associated lead and copper screening values developed during the

laboratory validation phase of this project, were readily validated in the field. Naturally aged soils and sediments were collected from each site and tested in the laboratory with naive salamanders and frogs to determine if site-specific conditions ameliorate toxicity. The naturally aged soils and sediments from each site have been exposed to the elements for months and years which may modify the bioavailability of certain compounds over time. At some sites where the chemicals have weathered for decades, the chemicals are held tightly by the soil and are unavailable for transport through the soil or into biological organisms (Loehr, 1996). The bioavailability of metals in soil will be influenced by the species present, particle size, and whether the metals have been encapsulated or coated by other mineral phases (Chaney, et al., 1989). The naïve test organisms have not previously been exposed to copper or lead so they are not expected to exhibit any natural resistance to the potential toxicity of these compounds. The test species used in the toxicity tests were the tadpoles of the Northern leopard frog (*R. pipiens*) for the sediment exposure assay and the adult red-backed salamander (*P. cinereus*) for the soil exposure assay.

The soil and sediments collected from the two demonstration sites were expected to contain a range of copper and lead concentrations as a result of current (Travis AFB) and historic (APG) firing range activities.

Three samples from each site (two impacted samples and one reference sample) were tested with the soil exposure protocol. Four impacted samples and a reference sample from each site were tested with the sediment exposure protocol. To provide a better distribution of lead concentrations in the tested sediments, three dilutions of the highest tested concentration were generated and included in the sediment exposure testing program. The sediment exposure tests also included a laboratory control sediment treatment with each set of tests.

Data from toxicity studies were statistically evaluated by comparing responses of individual soils/sediments to the laboratory control (standard, clean sediment that will not produce adverse effects) and/or through comparison to a reference soil/sediment collected concurrently with other samples.

Both lethal (mortality) and sub-lethal (growth) endpoints were measured in each toxicity study. Each data type was analyzed independently. Because sub-lethal measurements of surviving organisms can be skewed by a significant reduction in sample size, treatments that demonstrated significant mortality were excluded from sub-lethal analyses of growth endpoints.

Before conducting tests to identify statistical differences, suitability of the data for parametric analyses was evaluated through normality and equality of variance tests. Based on the results of the tests for normality of distribution and homogeneity of sample variances, data sets were evaluated using the appropriate parametric or non-parametric Analysis of Variance (ANOVA) statistic. Pair-wise comparisons were conducted to determine statistical differences between tested samples and reference or control results. Statistical difference was evaluated at α=0.05. The alpha level represents the probability (5%) of committing a Type I statistical error, that is, finding a significant difference when, in fact, one does not exist. Reducing the alpha level will reduce the probability of committing a Type I error, but will also increase the probability of committing a Type II error, or not finding a statistical difference when, in fact, one does exist. For that reason, 0.05 is generally accepted as a reasonable level for most analyses.

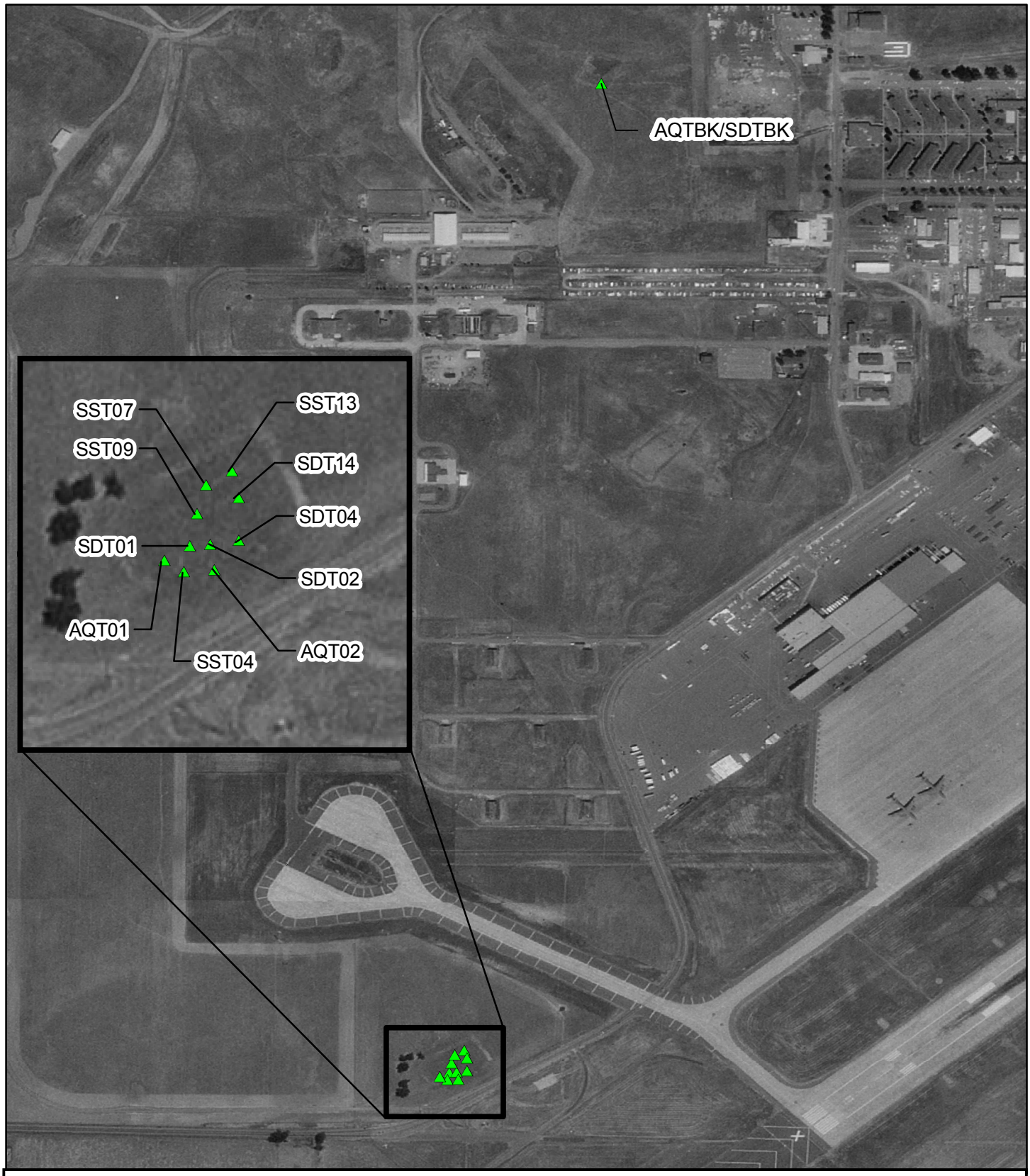
3.5.7 Sampling Plan

The sampling plan for both demonstration sites used the results of XRF surveys to identify appropriate sampling locations. Sampling was conducted in accordance with the Quality Assurance Project Plan presented in Appendix C of the *Field Demonstration Plan* (NAVFAC, 2007a) and the completed document is on file at ENSR and available upon request. Additional detail for each demonstration site is provided in the following subsections.

3.5.7.1 Travis Air Force Base (Travis AFB)

As described in Section 3.5.1, the results of the analytical samples collected during the field reconnaissance effort (Figure 3-3) were used to direct the XRF survey conducted as part of the field sampling event on March 27, 2006. The XRF survey included analysis of 26 samples for 22 metals.

Based on the results of the XRF survey, soil and sediment samples were collected from nine locations within the vernal pool study area and one reference location (Figure 3-4). Based on the moisture content of the material, five samples, including the reference, were identified as sediment and the remaining four were identified as soils. Table 3-6 presents the copper and lead data from the XRF survey and the confirmation analyses conducted at Paragon Analytical, Inc (Fort Collins, CO) for the nine samples collected as part of the Travis AFB field demonstration.



CA Digital Orthophoto Quads, 2002
Elmira, CA
Approximate Sampling Locations,
3/28/06

N

0 445 890 1,780 Feet

Travis Air Force Base Sampling Locations

SCALE	DATE	PROJECT NO.
1:12000	10/07	09070-056

ENSR | AECOM

Figure Number

3-4

J:\Govt\Projects\ESTCP 2005 Amphibians\Technology Demonstration Effort and Results\TechDemoField\TravisAFB\Travis\GIS\samplingTravis.mxd

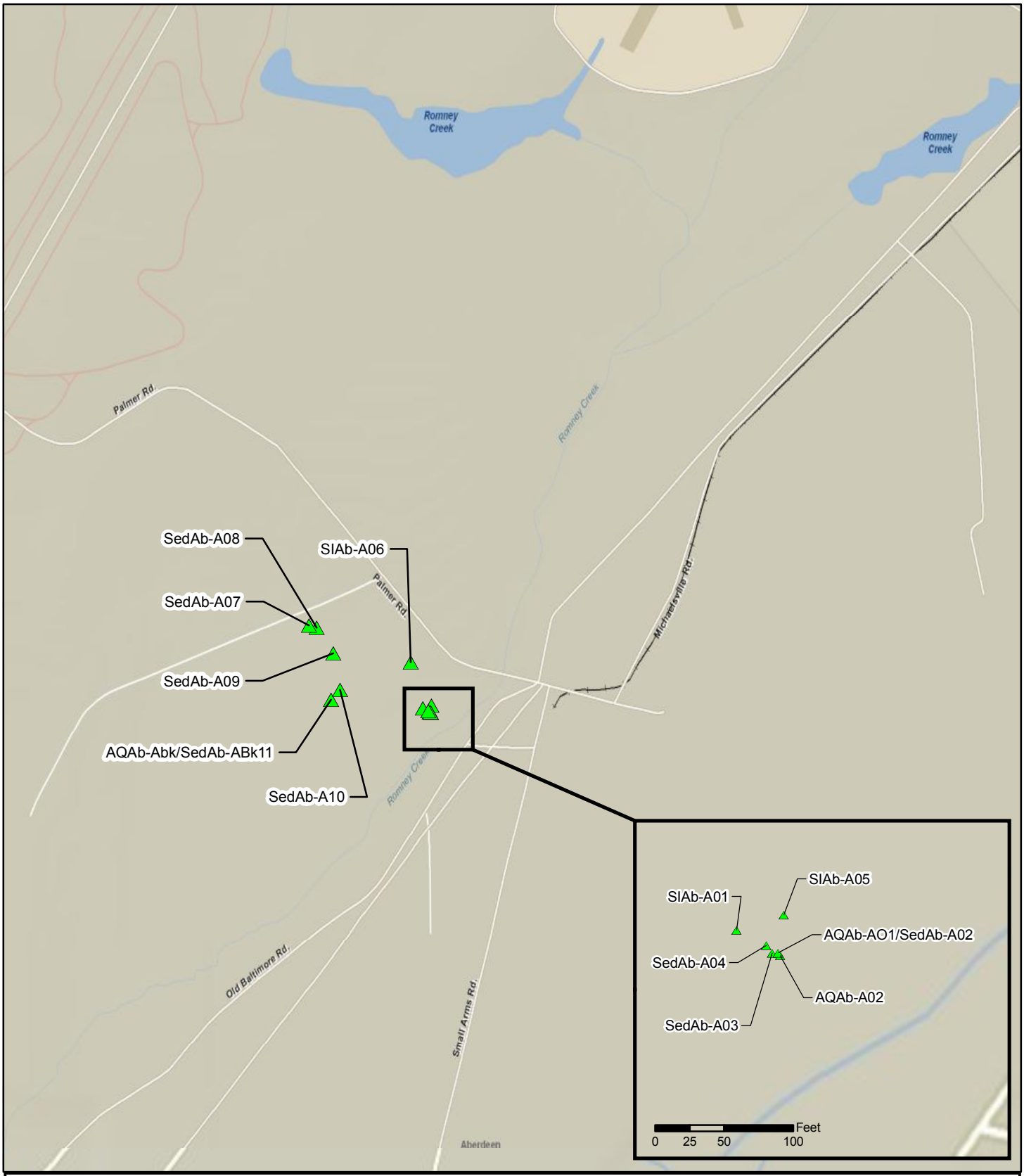
All sediment and soil samples were analyzed for copper, lead, TOC, grain size, CEC, simultaneously extracted metals (SEM), and acid volatile sulfides (AVS) (Table 3-6). Three samples were also analyzed for a full suite of 23 metals, 21 pesticides, and polychlorinated biphenyls (PCB) as Aroclors. The background sample was also analyzed for 17 PAHs. PAHs, pesticides, and PCBs were not detected in any samples. Metals, other than lead, were present at low levels and concentrations were similar between vernal pool samples and the reference location.

Surface water samples were collected from 2 locations within the vernal pool (center of pool and at the outlet) and at the reference location. Samples were analyzed for 23 total recoverable metals, total hardness, TOC, dissolved organic carbon (DOC), and dissolved phase copper, lead, and hardness. Fourteen metals were not detected in any of the vernal pool surface water samples. The detected metals (barium, calcium, iron, lead, magnesium, manganese, potassium, selenium, and sodium) were detected at low levels that would not be expected to be acutely toxic to aquatic receptors. Copper, lead, hardness, and organic carbon results are presented in Table 3-7.

3.5.7.2 Aberdeen Proving Ground (APG)

The APG study area has been the site of historic XRF data collection efforts as part of previous site investigation activities. These existing data were used to focus the XRF survey conducted just prior to the field sampling effort on specific areas of interest with lead concentration gradients desired for conducting the toxicity tests (see Section 3.5.1 for discussion of the XRF survey). The APG study area was also limited to the area cleared of UXO by Army personnel.

Based on the results of the XRF survey, samples were collected from ten locations within the palustrine wetland study area and one reference location (Figure 3-5). Based on the moisture content of the material, eight samples, including the reference, were identified as sediment and the remaining three were identified as soils. Table 3-8 presents the copper and lead data from the XRF survey and the confirmation analyses.



ArcGIS Online Imagery Service 2008
Aberdeen, MD
GPS Coordinates of Sampling Locations, 4/12/06

N

Feet
0 1,000 2,000 4,000

Aberdeen Proving Ground Sampling Locations

SCALE	DATE	PROJECT NO.
1:21,141	12/08	09070-056

ENSR | AECOM

Figure Number

3-5

J:\Govt\Projects\ESTCP 2005 Amphibians\Technology Demonstration Effort and Results\TechDemoField\Aberdeen\Aberdeen\GIS\Sampling.mxd

Table 3–6 Analytical Results for Sediment and Soil Collected from Travis Air Force Base

Location ID	Sample Matrix	Chemical Concentration (mg/kg)				TOC (%)	[Sum SEM - AVS]/f_{oc} (umol/g_{oc})	Bulk Density (lb/ft3)	Cation Exchange Capacity (meq/100g)	Grain Size (%)		
		XRF Survey [1]		Analytical Laboratory [2]								
		Copper	Lead	Copper	Lead					Gravel	Sand	Silt & Clay
Travis Air Force Base												
SDTBK	sediment	<16	20	15	14	1.5	-65.4	110	9.2	0	47.3	52.7
SST04	soil	<17	464	16 [3]	935 [3]	1.75 [3]	***325*** [3]	112	1.5	0.8	46.7	52.5
SDT04	sediment	no data	no data	21	1500	1.8	NC	106	13.4	0.4	45	54.6
SST09	soil	<18	1802	16	1600	1.6	***409***	116	1.3	0.9	48.9	50.2
SST07	soil	<20	4025	17	2000	1.6	***466***	110	8.0	0.3	48.1	51.6
SDT01	sediment	<15	910	20	2100	2	***452***	100	2.1	0	47.5	52.5
SDT02	sediment	<18	1268	20	2100	1.6	NC	114	1.5	0	47.4	52.6
SDT14	sediment	<134	1459	17	2500	1.9	***987***	110	11.9	0	45.3	54.7
SST13	soil	<120	2928	19	4200	2.3	***726***	106	1.8	0.5	47.8	51.7

BK in location ID identifies background reference location; all others are site locations.

1 - Samples analyzed by XRF survey prior to sediment and soil sampling. Values are not corrected for percent moisture (corrected values would be approximately 25% lower; assuming percent solid is approximately 75%).

2 - Samples analyzed by analytical laboratory directly following field sampling effort. Values reported on a dry weight basis.

3 - Average of parent and duplicate sample results.

One-half detection limit used in calculation for non-detect SEM analytes.

Shading and bold text indicates [Sum SEM-AVS]/f_{oc} > 130 umol/g_{oc}

- Values < 130 umol/g_{oc} are presumed to be "not likely" to be toxic (USEPA, 2005a).

f_{oc} - fraction organic carbon

g_{oc} - gram organic carbon

NC - Not calculated; Calculations not completed if AVS was not detected.

Table 3–7 Analytical Results for Surface Water Collected from Travis Air Force Base

	Chemical Concentration (mg/L)							
	Copper		Lead		Hardness		Organic Carbon	
Location ID	Dissolved	Total	Dissolved	Total	Dissolved	Total	Dissolved	Total
Travis Air Force Base								
AQT01 [1]	0.01 U	0.01 U	0.012	0.017	47	46.5	11	12.5
AQT02	0.01 U	0.01 U	0.0078	0.016	50	50	12	13
AQTBK	0.01 U	0.01 U	0.003 U	0.003 U	17	22	15	22

BK in location ID identifies background reference location; all others are site locations.

1 - Average of parent and duplicate sample results for total copper, lead, hardness, and organic carbon.

U – Not detected at or above the stated detection limit.

All sediment and soil samples were analyzed for copper, lead, TOC, grain size, CEC, SEM, and AVS (Table 3-8). Two samples and a field duplicate were also analyzed for a full suite of 23 metals, 21 pesticides, and PCB Aroclors. Pesticides and PCBs were not detected in any samples, with one exception. 4,4-DDE was detected in one sample at 6.2 micrograms per kilogram (ug/kg). This level is between the threshold effect concentration (TEC; 3.2 ug/kg) and the probable effect concentration (PEC; 31.3 ug/kg) discussed in MacDonald et al. (2000), indicating that the possibility of toxicity to benthic receptors could not be excluded in this sample. Most metals, other than lead, were present at low levels. Copper was present at higher concentrations than were observed at Travis AFB. Copper concentrations increased with increasing lead concentrations to a maximum of 1,150 mg/kg copper.

Surface water samples were collected from two locations and a field duplicate within the wetland and at the reference location. Samples were analyzed for 23 total recoverable and dissolved phase metals, total and dissolved hardness, TOC, and DOC. Fourteen metals were not detected in any of the surface water samples. The detected metals (calcium, copper, iron, lead, magnesium, manganese, potassium, sodium, and zinc) were detected at low levels that would not be expected to be acutely toxic to aquatic receptors. Copper, lead, hardness and organic carbon results are presented in Table 3-9.

3.5.7.3 Sample Collection

As described in Section 3.5.7, surface water, sediment, and hydric soil samples were collected from both Travis AFB and APG for chemical and toxicological analyses. During sample handling it was critical that cross-contamination between samples, or contamination from extraneous sources, not occur. Sample handling tools were constructed of inert materials wherever possible (stainless steel, polypropylene, or Teflon, as appropriate) and were decontaminated between sampling locations.

Decontamination consisted of a tap water rinse to remove gross contamination (if needed), followed by a non-phosphate detergent (e.g., Alconox) water rinse, a rinse with deionized water, and followed by another deionized water rinse. If equipment was to be stored or transported, it was wrapped in aluminum foil after air-drying. Water generated during decontamination of sampling equipment was containerized and disposed of according to DoD facility instructions.

Surface water samples were collected from mid-depth at selected sediment sampling locations prior to the collection of the sediment sample. All sediment samples were collected from relatively shallow locations so a boat was not required and samplers could wade in to the stations. Sediment was generally collected using stainless steel trowels and spoons. The sample was collected from the top 6 inches of sediment, with as little disturbance as possible. Soil samples were collected from the surficial 6 inches also using stainless steel trowels and spoons.

Soil and sediment samples from each sampling location were composited in a large stainless steel bowl prior to sub-sampling for chemical and toxicological analyses. To allow for accidental loss, spillage, analytical chemistry, or test reruns, a minimum of two gallons of each sediment and soil sample was collected from each location. Samples were cooled to 4°C before shipping and when not being used.

Table 3–8 Analytical Results for Sediment and Soil Collected from Aberdeen Proving Ground

Location ID	Sample Matrix	Chemical Concentration (mg/kg)			TOC (%)	[Sum SEM - AVS]/f_{oc} (umol/g_{oc})	Bulk Density (lb/ft3)	Cation Exchange Capacity (meq/100g)	Grain Size (%)		
		XRF Survey [1]	Analytical Laboratory [2]								
		Lead	Copper	Lead					Gravel	Sand	Silt & Clay
Aberdeen Proving Ground											
SedAb-ABk11	sediment	33	8.4	34	0.46	NC	123	7.3	0.2	46.2	53.6
SlAb-A06	soil	38	17	35	1.3	NC	119	9	0	26.1	73.9
SedAb-A08	sediment	343	21	200	0.36	-40.3	101	3.1	3.3	87.7	9
SlAb-A05	soil	1542	55	260	0.41 [3]	***371*** [4]	102	5.1	0	45.2	54.8
SedAb-A10	sediment	168	27	310	0.62	NC	112	13.8	0.3	14.2	85.5
SedAb-A07	sediment	229	34	460	1.9	49	106	10.9	3.9	36.4	59.7
SedAb-02B	sediment	1200	120 [3]	1275 [3]	0.55	***581***	128	1.6	2.3	59.5	38.2
SedAb-A04	sediment	4001	120	850	0.57	NC	117	4.3	2.9	47.7	49.4
SedAb-A09	sediment	741	73	870	0.11	***3336***	121	0.6	1.6	82.9	15.5
SlAb-A01	soil	12387	700	9900	0.088	NC	125	0.9	3.9	76.2	19.9
SedAb-A3A	sediment	2000	1150 [3]	17500 [3]	2.0 [3]	NC [3]	100	5.8	0	23.3	76.7

Bk in location ID identifies background reference location; all others are site locations.

1 - Samples analyzed by XRF survey prior to sediment and soil sampling. Values are not corrected for percent moisture (corrected values would be approximately 25% lower; assuming percent solid is approximately 75%).

2 - Samples analyzed by analytical laboratory directly following field sampling effort. Values reported on a dry weight basis.

3 - Average of parent and duplicate sample results.

4 -AVS was not detected in duplicate sample so only results of parent sample are presented.

Shading and bold text indicates [Sum SEM-AVS]/f_{oc} > 130 umol/g_{oc}

- Values < 130 umol/g_{oc} are presumed to be "not likely" to be toxic (USEPA, 2005a).

f_{oc} - fraction organic carbon

g_{oc} - gram organic carbon

NC - Not calculated; Calculations not completed if AVS was not detected.

NS - Not sampled.

One-half detection limit used in calculation for non-detect SEM analytes.

Table 3–9 Analytical Results for Surface Water Collected from Aberdeen Proving Ground

	Chemical Concentration (mg/L)							
	Copper		Lead		Hardness		Organic Carbon	
Location ID	Dissolved	Total	Dissolved	Total	Dissolved	Total	Dissolved	Total
Aberdeen Proving Ground								
AQAb-Abk	0.01 U	0.01 U	0.003 U	0.004	13	13	12	14
AQAb-A01	0.04	0.045	0.027	0.063	15	15	13	16
AQAb-A02 [1]	0.038	0.0555	0.0395	0.14	14	15.5	12	15.5

Bk in location ID identifies background reference location; all others are site locations.

1 - Average of parent and duplicate sample results for total copper, lead, hardness, and organic carbon.

U – Not detected at or above the stated detection limit.

Samples were accompanied by a properly completed chain-of-custody form. This record documents the transfer of custody of samples from the sampler, to another person, and to the permanent laboratory. Shipping containers were secured with strapping tape and sealed with custody seals. Samples were shipped daily from the field to the laboratory using an overnight courier.

Holding times are listed in Table 3-10. Holding times were met for metal, PCB, pesticide, and PAH analyses in soil and sediment and for all surface water analytes. Holding times were not met for the TOC or SEM and AVS analyses. Impacts to the data due to exceeded holding times are discussed in Section 3.5.7.6.

Table 3–10 Sample Holding Time Requirements

Parameters	**Holding Time[1]**
Sediment/Soil	
Metals	28 days for Hg; 6 months for others
TOC	14 days
Pesticides	14 days to extraction; 40 days from extraction to analysis
PCBs	14 days to extraction; 40 days from extraction to analysis
PAHs	14 days to extraction; 40 days from extraction to analysis
SEM/AVS	14 days
Grain size	None
Cation exchange capacity	None
Bulk density	None
Surface Water	
Metals (dissolved)	28 days for Hg; 6 months for others
Metals (total recoverable) and hardness	28 days for Hg; 6 months for others
DOC	28 days
TOC	28 days
1 - Holding time begins from date of sample collection.	

QC samples collected during the field sampling effort at each site included an equipment rinsate blank, a field duplicate for each medium, and MS/MSDs as appropriate for the parameter and media sampled.

Figure 3-6 Sediment Collection at Travis Air Force

Figure 3-7 Sediment Sampling Location at Travis Air Force Base

3.5.7.4 Sample Analysis

Sediment and soil samples were shipped on ice under chain-of-custody directly to the chemistry and toxicity testing laboratories. All sediment and soil samples were analyzed for copper, lead, TOC, grain size, CEC, SEM, and AVS following the field sampling effort. A subset of sediment and soil samples were also analyzed for a full suite of 23 metals, 21 pesticides, 17 PAHs, and 7 PCB Aroclors. Surface water samples were analyzed for 23 total recoverable and dissolved phase metals, total and dissolved hardness, TOC, and DOC.

Copper and lead were considered to be the primary chemicals of concern in the sediment and soil. Detection limits ranged from 1.0 to 1.8 mg/kg for copper and 0.31 to 47 mg/kg for lead in sediment and soil. The elevated detection limits for lead were reported for samples requiring dilutions due to concentrations that exceeded the calibration range.

Samples selected for toxicity testing are presented in Table 3-11. Copper and lead concentrations in soil and sediment were reviewed in order to select a wide range of concentrations. The sediment toxicity testing was conducted at ENSR's Fort Collins Environmental Toxicology Laboratory (FCETL) in Fort Collins, Colorado and the soil toxicity testing was conducted at USACHPPM's APG , Maryland laboratory. Toxicity testing was conducted according to the quality assurance/quality control (QA/QC) plans in place at the ENSR and the USACHPPM toxicity laboratories and the protocols presented in Appendix A.

Table 3–11 Samples Selected For Toxicity Testing

Location ID	Sample Matrix	Chemical Concentration [1] (mg/kg)		Samples Selected For Toxicity Testing	
		Copper	Lead	Salamander [Soil Exposure]	Tadpole [Sediment Exposure]
Travis Air Force Base					
SDTBK	sediment	15	14	x	x
SST04 [3]	soil	16	935		
SDT04	sediment	21	1500		x [2]
SST09	soil	16	1600	x	
SST07	soil	17	2000		x
SDT01	sediment	20	2100		
SDT02	sediment	20	2100		
SDT14	sediment	17	2500		x
SST13	soil	19	4200	x	x
Aberdeen Proving Ground					
SedAb-ABk11	sediment	8	34		x
SlAb-A06	soil	17	35	x	
SedAb-A08	sediment	21	200		x
SlAb-A05	soil	55	260	x	
SedAb-A10	sediment	27	310		
SedAb-A07	sediment	34	460		x
SedAb-02B [3]	sediment	120	1275		
SedAb-A04	sediment	120	850		x
SedAb-A09	sediment	73	870		
SlAb-A01	soil	700	9900	x	
SedAb-A3A [3]	sediment	1150	17500		x [2]

BK or Bk in location ID identifies background reference location; all others are site locations.

1 - Samples analyzed by analytical laboratory following field sampling effort.

2 - Three dilutions from this sample were generated and tested in order to achieve a better distribution of lead concentrations.

3 - Average of parent sample and duplicate results for copper and lead.

Due to delays in starting the sediment toxicity tests associated with test organism availability, samples selected for the sediment toxicity testing were submitted for an additional set of copper and lead analyses prior to test set-up. Lead in soil was analyzed by an in-house USACHPPM laboratory on Days 0, 14, and 28 of the soil exposure test. These additional sets of analytical data are presented in Section 4 with the results of the sediment and soil toxicity tests.

3.5.7.5 Experimental Controls

At least one negative control treatment was included in the sediment and soil tests. Negative controls represent sediment and soil without significant levels of copper or lead. This may be accomplished through the use of a "clean" sample provided by the laboratory or through the collection of a reference control soil or sediment that represents a similar habitat type to the study area, but does not contain significant levels of copper or lead.

Reference samples were collected in both soil and sediment at each of the demonstration sites. Reference locations with similar physical characteristics (e.g., organic carbon, grain size) to the tested samples were selected to avoid the impact of these characteristics on the interpretation of the test results. The reference soil samples served as the negative controls in the soil tests. The sediment tests included the sediment reference samples as well as laboratory control samples as the negative controls.

3.5.7.6 Data Quality Parameters

The *Field Demonstration Plan* (NAVFAC, 2007a) specified collection and handling procedures designed to ensure the representativeness and integrity of the samples. The analytical program was designed to generate definitive data of sufficient quality and sensitivity to meet the project objectives.

Laboratory QA/QC measures were performed by the analytical laboratories to ensure that all environmental efforts to produce the data are technically sound and legally defensible. The quality assurance measures include standard operating procedures, applicable certifications, training programs, internal audits, and internal QC checks. Internal QC checks differ slightly for each individual procedure but in general include the following:

- Method blanks – used to define the level of laboratory background and reagent contamination,
- Laboratory control spikes (LCSs) – provide information on method accuracy and laboratory performance,
- Matrix spikes – determine accuracy of the method for the matrix,
- Duplicate samples – used to demonstrate acceptable method precision by the laboratory at the time of analysis,
- Surrogate spikes – used to detect problems in sample preparation procedures.

The laboratory SOPs for each analysis define the type, frequency, and corrective action for the applicable QC checks.

For the sediment toxicity testing laboratory, laboratory controls were tested with each set of samples and reference toxicant tests were completed within the acceptable results range.

Measures to ensure representativeness, completeness, comparability, accuracy and precision of the data are discussed below and in the Quality Assurance Project Plan (QAPP; Appendix B).

Representativeness – Representativeness expresses the degree to which data accurately and precisely represent a characteristic of a population, parameter variations at a sampling point, a process condition, or an environmental condition within a defined spatial and/or temporal boundary. Therefore, representativeness of the field data is highly dependent upon the proper design of the sampling program, and was achieved through strict adherence to the Field Demonstration Plan (NAVFAC, 2007a). Representativeness of the laboratory data was achieved by strict adherence to analytical SOPs and conformance to the majority of sample holding times. Other than holding time nonconformances for TOC, SEM, and AVS analyses, no significant deviations were noted. Exceeding the holding time for TOC analysis would not be expected to significantly impact the sample results. Exceeding the holding time for AVS analyses may result in an under-estimation of the sulfide levels as they dissipate over time. It is possible that the SEM mercury results are biased low due to the exceeded holding time.

Completeness – Completeness is a measure of the amount of valid data obtained from a measurement system compared to the amount that was expected to be obtained under normal conditions. "Normal conditions" are defined as the conditions expected if the sampling plan was implemented as planned.

Field sampling completeness was 100%. Samples from all proposed stations were collected and submitted for the analyses in the QAPP.

Laboratory completeness is a measure of the amount of valid measurements obtained from all the measurements taken in the project. All analyses were successfully performed by the laboratory for the requested parameters. However, soil and sediment pH measurements were inadvertently not performed.

Comparability – Comparability expresses the confidence with which one data set can be compared to another. Analytical data from this program are considered to be comparable since similar sampling and analytical methods were used at each demonstration site as documented in the Field Demonstration Plan (NAVFAC, 2007a).

Accuracy – Accuracy is the degree of agreement between the observed value and an accepted reference or true value. Accuracy in the field was assessed through the use of equipment rinsate blanks and through the adherence to sample handling, preservation, and holding time requirements. The objective for equipment rinsate blanks is that no target analytes are present above the reporting limit. This objective was achieved for the equipment rinsate blanks collected at each demonstration site except for one equipment rinsate blank associated with the soils and sediments collected at APG (see Appendix C). Zinc was detected in the equipment rinsate blank at a concentration of 0.040 mg/L (2x the detection limit). This level of zinc contamination is insignificant compared to the zinc concentrations in the samples and does not impact sample results.

Laboratory method blanks were free of contamination for all parameters except SEM zinc and mercury. Low level concentrations of these analytes reported in the samples analyzed for SEM are likely to be biased high.

The impacts to sample data due to exceeded holding times are presented above as part of the discussion on representativeness.

Laboratory accuracy was assessed through the analysis of matrix spike/matrix spike duplicates (MS/MSDs), LCSs, and surrogate compounds, and the subsequent determination of the recoveries of the spiked analytes.

Due to a laboratory oversight, MS/MSD analyses were not performed for TOC. However, LCSs were analyzed with each batch of samples. Acceptable recoveries were obtained for the LCSs demonstrating acceptable laboratory performance of the method.

In general, MS/MSD, LCS, and surrogate recoveries fell within the laboratory control limits. MS recoveries for AVS were very low, most likely due to difficulties with the sample matrix. AVS results should be considered to have a very low bias.

Precision - Precision is a measure of the degree to which two or more measurements are in agreement. Precision was measured through the calculation of relative percent difference (RPD). The objectives for field precision RPDs are 30% RPD for aqueous samples and 50% RPD for solid samples. Overall, the objectives for field precision were met with only a few exceptions. The RPDs for total lead in the surface water field duplicate pair collected at APG and for total manganese in the surface water field duplicate pair collected at Travis AFB exceeded the RPD criterion. The RPDs for antimony and lead exceeded the criterion for the sediment field duplicate pair collected at APG. The RPD for SEM lead exceeded the criterion in the soil field duplicate pair collected at APG. High RPDs in solid samples are most likely due to sample non-homogeneity.

Precision in the laboratory is assessed through the calculation of RPD for duplicate samples, either as MS/MSDs or as laboratory duplicates, depending on the method. The laboratory utilized current in-house control limits at the time of analysis for assessing precision. In general, RPDs for laboratory duplicates and MS/MSDs met the acceptance criteria. Sample nonhomogeneity was likely the source of the high RPDs observed for selected metals in the solid samples.

3.5.7.7 Data Quality Indicators

Statistical analyses were used to identify significant differences in the soil and sediment toxicity tests. Organism responses for both lethal and sub-lethal endpoints were compared against the associated laboratory control or reference station results.

Equations used to evaluate analytical data quality are presented in the QAPP (Appendix B).

3.5.7.8 Calibration Procedures, Quality Control Checks, and Corrective Action

Field instruments used during the sampling events included a GPS unit, an XRF unit, and YSI multi-parameter meter. All field instruments were free from obvious defects, damage, and contamination and were properly functioning during the field events. Daily operational checks and calibrations were conducted. In general, field instruments were calibrated prior to daily use, and were checked after every 15 samples and at the end of the day. Calibration procedures were consistent with the manufacturer's recommendations.

Routine testing and preventive maintenance is performed by the analytical laboratories as part of their in-house QA programs. Calibration procedures for laboratory instruments generally consisted of initial calibrations, initial calibration verifications, and continuing calibration verification in accordance with the SOP for each analysis.

For the toxicity laboratories, the performance of test organisms in the laboratory control and/or the reference samples is used to determine test acceptability. As described in Section 3.5.7.5, these negative controls represent sediment and soil without significant levels of copper or lead and these samples may be used to evaluate the health of the test organisms and the test conditions (e.g., lighting, temperature). If survival in the control treatment (or the reference sample in the soil exposure assay) is less than 80%, then the test data should be carefully examined to determine if it is acceptable. Survival in the negative controls was >80% in the soil and sediment tests; therefore all tests were considered to be acceptable

Reference toxicant tests were also conducted to evaluate the health of the tadpoles used in the sediment exposure test. The response of test organisms in the reference toxicant tests and the negative controls indicated that the health of the test organisms was not impaired prior to exposure to the test sediments.

3.5.8 Demobilization

Following the completion of field surveys and sample collection at the selected sites, samples were shipped to the appropriate analytical laboratory or toxicity testing laboratory for analysis. Sampling and homogenization equipment was washed and decontaminated between samples and prior to demobilization. Sampling equipment was shipped back to the appropriate point of origin.

3.6 Selection of Analytical/Testing Methods

Protocols for conducting the sediment and soil bioassays are provided in Appendix A. These protocols were developed under previous DoD-funded projects. The laboratory validation phase of this project was designed to finalize the protocols and the associated protocols (NAVFAC, 2007b). Analytical methods are described in the QAPP presented in Appendix B.

3.7 Selection of Analytical/Testing Laboratory

The sediment and soil toxicity testing was conducted at the FCETL and the USACHPPM toxicity laboratories, respectively. The sediment and soil exposure protocols were developed and validated at these laboratories and both facilities have been involved in conducting similar types of tests for many years.

Chemical analyses were conducted by the following laboratories:

- Paragon Analytics of Fort Collins, Colorado – metals, pesticides, PCBs, PAHs in soil, sediment and water
- Mitkem Corporation of Warwick, Rhode Island – TOC
- STL-Burlington of Colchester, Vermont – SEM and AVS

- GeoTesting Express of Boxborough, Massachusetts – grain size, bulk density, and CEC analyses

The soil toxicity testing protocol requires some special expertise in animal handling (e.g., for test termination) and in tissue processing and analysis (e.g., for histopathological and blood chemistry endpoints). The sediment testing protocol should not require expertise beyond what would be typical for a laboratory conducting sediment toxicity tests with invertebrates or fish. The chemical analyses conducted to characterize the soil and sediment are typical for site investigations and do not require any special processing or expertise.

4.0 Performance Assessment

4.1 Performance Criteria

Performance criteria for the field demonstration are presented in Table 4-1 and are based upon the performance objectives presented in Table 3-1. These criteria were originally presented in the *Field Demonstration Plan* (NAVFAC, 2007a). The types of performance objectives and criteria established for typical remediation-related ESTCP projects (e.g., end-point criteria, remediation time, and analytical sensitivity) are indirectly associated with the ecological risk and toxicity based performance objectives and criteria developed for this project. The success of the performance of the innovative technology was determined based on whether or not the soil and sediment exposure protocols were able to correlate an amphibian response with contaminant concentrations and whether the protocols could be broadly applied at sites requiring risk assessment characterization for amphibians.

Table 4–1 Performance Criteria

Performance Criteria	Description	Primary or Secondary
Sediment protocol is applicable to evaluating copper and lead in palustrine wetlands	Describe whether or not there is a statistical relationship between contaminant concentrations and the results of the assay.	Primary
Soil protocol is applicable to evaluating copper and lead in forested uplands	Describe whether or not there is a statistical relationship between contaminant concentrations and the results of the assay.	Primary
Collection and biological evaluation of native salamanders is applicable for evaluating potential impacts due to metals	Describe whether or not there is a relationship between contaminant concentrations and the blood and/or histological evaluation.	Secondary
Regulatory acceptance of toxicity test protocols	Describe interaction with regulatory agencies regarding amphibian ERA results	Primary
Versatility of the overall ERA protocol	Describe whether or not ERA protocol was applicable at both field demonstration sites	Primary
Technology transferred to other potential end-users	Identify and describe presentations of the technology at conferences or in journals.	Secondary
Sediment Exposure Protocol - Sediment toxicity test is valid and acceptable	Describe whether test acceptability criteria were met	Primary
Sediment Exposure Protocol - Lethal endpoint indicates toxicity or lack of toxicity	Describe whether statistical differences were observed between tested samples and the control or reference stations	Primary

Table 4–1 Performance Criteria (continued)

Performance Criteria	Description	Primary or Secondary
Sediment Exposure Protocol - Sub-lethal endpoints indicate toxicity or lack of toxicity	Describe whether statistical differences were observed between tested samples and the control or reference stations	Primary
Soil Exposure Protocol - Soil toxicity test is valid and acceptable	Describe whether test acceptability criteria were met	Primary
Soil Exposure Protocol - Lethal endpoint indicates toxicity or lack of toxicity	Describe whether statistical differences were observed between tested samples and the control or reference stations	Primary
Soil Exposure Protocol - Growth endpoints indicate toxicity or lack of toxicity	Describe whether statistical differences were observed between tested samples and the control or reference stations	Primary
Soil Exposure Protocol - Blood parameters indicate toxicity or lack of toxicity	Describe whether statistical differences were observed between tested samples and the control or reference stations	Primary

This study was conducted consistent with the standards found in Title 40 Code of Federal Regulations (CFR), Part 792, Good Laboratory Practices and through an approved protocol with the Institutional Animal Care and Use Committee (IACUC). The investigators and technicians adhered to the following guidelines: the Public Health Service Policy on Humane Care and Use of Laboratory Animals, "U.S. Government Principles for the Utilization and Care of Vertebrate Animals Used in Testing, Research, and Training," and the Animal Welfare Act.

Although not identified in the *Field Demonstration Plan* (NAVFAC, 2007a) as a formal performance criteria, in order to comply with DoD requirements the two laboratories conducting the toxicity tests were required to have appropriate Animal Use Protocols in-place. This Protocol ensures that all vertebrate animals are treated humanely and do not endure any unnecessary pain.

The CHPPM laboratory conducting the soil exposure testing had established the appropriate Animal Use Protocols prior to the laboratory validation phase of testing. The FCETL did not already have an existing Animal Use Protocol so a project-specific IACUC was convened to oversee and evaluate the animal care program and ensure that treatment of animals at the laboratory was in compliance with applicable regulations and laws (e.g., animal welfare regulations).

The project-specific IACUC toured the laboratory facility, reviewed documentation, questioned researchers, and provided conditional approval of the testing protocol in April 2006. To achieve final approval from the IACUC, the laboratory modified their animal care standard operating procedure (SOP) and generated a project-specific Animal Care and Use Questionnaire for conducting the sediment protocol and submitted both documents for review in August 2006. Final approval from the IACUC was received in September 2006. The final approval memo, the SOP

and the questionnaire are provided in Appendix A of the *Test Refinement Interim Report* (NAVFAC, 2007b).

4.2 Performance Confirmation Methods

Adherence to the data collection methods and analyses presented in the *Field Demonstration Plan* (NAVFAC, 2007a) ensured that reliable data was collected. Data quality was assessed through the use of duplicate analytical samples, MS/MSD analyses, and the use of negative controls in the toxicity tests. Sufficient data were collected to evaluate the performance criteria listed in Table 4-1. Table 4-2 presents the expected performance metric for each of the performance criteria, the method that was used to confirm performance, and the actual performance noted during the demonstration.

The primary measurement for determining the effectiveness of the demonstration was whether or not there was a relationship between the concentrations of copper and/or lead in the soil and sediment and the results of the associated toxicity tests. Statistical methods were used to determine whether or not test organism responses in the tested samples containing elevated levels of lead were different from responses in laboratory controls or reference sample containing much lower levels of lead.

4.3 Data Analysis, Interpretation and Evaluation

As described in Section 3.5.6, statistics were used to evaluate whether or not toxic responses in tested soil or sediment at each demonstration site were significantly different from the laboratory control or reference stations. Following the statistical evaluation the analytical chemistry data were reviewed in order to identify media concentrations that may correlate with a toxic response. A lead concentration gradient was tested at both sites, allowing the development of LOAECs and NOAECs for both survival and sub-lethal endpoints. To derive these values, the survival or sub-lethal data for all stations at a site were ranked by the associated lead concentration with an indication of which samples were statistically toxic compared to the reference locations.

Some tested samples were identified as toxic compared to the reference while others were consistent with the reference results, indicating a non-toxic response. LOAECs and NOAECs were estimated by identifying the concentration of each analyte at the demarcation between toxic and non-toxic samples, as indicated by the statistical evaluation. The NOAEC represents the tested sample with the highest concentration of a constituent of potential concern (COPC) that was not significantly different from the control or reference station, whereas the LOAEC is the tested sample above which all concentrations were significantly different from the control or reference.

The results of the field demonstration tests were also evaluated relative to the screening values developed during the laboratory validation phase of testing (presented in Section 3.4) with lead-spiked soil and sediment.

Finally, the use of the amphibian ERA framework was evaluated at each site to determine whether it would be applicable for characterizing potential risks to amphibians at the two demonstration sites.

Table 4–2 Expected Performance and Performance Confirmation Methods

Performance Criteria	Expected Performance Metric	Performance Confirmation Method	Actual
Primary Criteria (Qualitative)			
Sediment protocol is applicable to evaluating copper and lead in palustrine wetlands	Correlation between sediment concentrations and lethal or sub-lethal results	Statistical evaluation to be conducted	Several samples with higher concentrations of lead from each demonstration site were statistically different from the reference samples
Soil protocol is applicable to evaluating copper and lead in forested uplands	Correlation between mesic soil concentrations and lethal or sub-lethal results	Statistical evaluation to be conducted	Several samples with higher concentrations of lead from each demonstration site were statistically different from the reference samples
Regulatory acceptance of toxicity test protocols	Results are accepted by agency as component of ERA	Study results submitted to regulatory agency as part of site assessment	Study results have not been submitted to agencies; no on-going investigations are being conducted at either demonstration site; however ASTM approval of the sediment protocol has been achieved
Versatility of the overall ERA protocol	ERA protocol applicable for various sites	Application of ERA protocol at both field demonstration sites	Tiered ERA protocol is appropriate for use at various sites

Table 4-2 Expected Performance and Performance Confirmation Methods (continued)

Performance Criteria	Expected Performance Metric	Performance Confirmation Method	Actual
Primary Criteria (Quantitative)			
Sediment Exposure Protocol - Sediment toxicity test is valid and acceptable	Mean survival in laboratory control is >80%	Laboratory controls evaluated at test termination	Laboratory control results met acceptability criteria
Sediment Exposure Protocol - Lethal endpoint indicates toxicity or lack of toxicity	Statistical difference between survival in control or reference samples and site samples	Statistical evaluation to be conducted	Statistical evaluation indicated significant mortality in some samples
Sediment Exposure Protocol - Sub-lethal endpoints indicate toxicity or lack of toxicity	Statistical difference between sub-lethal endpoints in control or reference samples and site samples[3]	Statistical evaluation to be conducted	Statistical evaluation indicated significant growth reduction (i.e., body width and length) in some samples
Soil Exposure Protocol - Soil toxicity test is valid and acceptable	Mean survival in laboratory control is >80%	Laboratory controls evaluated at test termination	Tests did not include laboratory control; survival was acceptable in reference samples
Soil Exposure Protocol - Lethal endpoint indicates toxicity or lack of toxicity	Statistical difference between survival in control and site samples	Statistical evaluation to be conducted	No lethal toxicity observed in any sample

[3] Sub-lethal endpoints may include growth, abnormalities, behavior, metamorphic stage, or other measurements.

Table 4-2 Expected Performance and Performance Confirmation Methods (continued)

Performance Criteria	Expected Performance Metric	Performance Confirmation Method	Actual
Primary Criteria (Quantitative)			
Soil Exposure Protocol - Growth endpoints indicate toxicity or lack of toxicity	Statistical difference between growth endpoints in control or reference samples and site samples	Statistical evaluation to be conducted	Statistical evaluation indicated significant growth reduction in several samples with higher lead concentrations relative to reference sample results
Soil Exposure Protocol - Blood parameters indicate toxicity or lack of toxicity	Statistical difference between blood parameters measured in control or reference samples and site samples	Statistical evaluation to be conducted	No statistical differences were observed in blood parameters in any samples
Secondary Criteria (Qualitative)			
Collection and biological evaluation of native salamanders is applicable for evaluating potential impacts due to metals	Correlation between mesic soil concentrations and histopathological evaluation	Statistical evaluation to be conducted	Native salamanders were not collected so criteria could not be evaluated
Technology transferred to other potential end-users	Presentation at conference or in journal	Results or protocols presented	Peer-reviewed article has been submitted to present soil exposure results. Sediment exposure protocol has been accepted as ASTM guide. Peer-reviewed articles to be prepared.

4.3.1 Soil Exposure Protocol

The 28-day soil exposure protocol was conducted in accordance with the general protocol described in Appendix A. The soils collected from each demonstration site were dried, pulverized, and sifted through 2 screens (Nalgene; 1- mm^2 and 0.5-mm^2 mesh) to homogenize the sample. Each treatment consisted of 10, individually housed salamanders (Figure 4-1). Each animal was placed into an individual Petri dish containing treatment-specific soil. Food consisted of potworms exposed to lead-contaminated soil from the same sample used to expose the salamanders. Animals were observed at least daily for signs of overt toxicity (e.g., lethargy, sensitivity to touch, abnormal behavior) and body weights were measured weekly.

Figure 4-1 Soil Exposure Protocol Test Set-up

On Day 28, surviving salamanders were euthanized using aqueous preparations of MS-222 followed by decapitation. The remaining head and body were preserved in 10% neutral buffered formalin. Cross sections of the head and body were then trimmed, embedded in paraffin, sectioned at 6 microns, stained with hematoxylin and eosin, and examined via routine light microscopy. The histologic sections were of adequate size and quality for the detection of treatment-related changes. Histologic observations and a record of tissues examined were entered into a computer-assisted data retrieval system (StarTox, Graham Laboratories, New Braunfels, TX) at the time of histologic examination. Growth, mortality, and health criteria (blood parameters, histological organ evaluation including quantification of liver melanomacrophages) results were incorporated into the dose response based screening values.

Soil samples were collected at the start (Day 0), mid-point (Day 14), and end of the assay (Day 28) to determine lead concentrations to which the salamanders were exposed during the assays. As described in Section 3.5.2, a protocol deviation resulted in half of the salamanders being exposed

for 29 days instead of 28 days. The necropsies were originally planned over the course of two days with the beginning exposures staggered accordingly; however all test organisms were necropsied on one day. Therefore, half the animals were exposed for an extra day.

4.3.1.1 Results of Field Demonstration

Three samples from each demonstration site were tested, with lead concentrations ranging from 11 mg/kg to nearly 17,000 mg/kg during the test (average of Day 0, Day 14, and Day 28 measurements). Copper levels in these samples measured just after the field effort (Tables 3-6 and 3-8) ranged up to 700 mg/kg; below the survival NOAEC observed during the laboratory validation testing (803 mg/kg; Table 3-3). A laboratory control treatment was not included in the test design so treatment results were compared against the results from the associated reference sample (SlAb-A06 was treated as the reference sample for the salamander tests on the APG samples).

Data from Travis AFB and APG were analyzed separately due to the differences in soil types (see Tables 4-3 and 4-4). The parameters evaluated included total erythrocyte counts, total leukocyte counts, hemoglobin, body weight, and percent change in body weight (calculated as the change from baseline body weight). Normality for all parameters was tested with Kolmogorov-Smirnov and transformed if necessary. Erythrocyte and leukocyte data were transformed with the natural log for analysis. One-way analysis of variance (ANOVA) was performed for all parameters using the lead concentration as the fixed factor. Repeated measures ANOVA was also performed on the weekly body weight measurements and percent change in growth. Tukey's post-hoc was performed if a significant difference ($p<0.05$) was found.

As indicated in Table 4-3, no mortality was observed in any treatments. Statistical evaluations indicated no significant differences among the Travis AFB samples.

The only difference observed among the APG samples was for percent change in body weight on Day 28 (Figure 4-2). Growth was significantly lower for SlAb-A01 (16,967 mg/kg) relative to the APG reference sample (SlAb-A06 with 28 mg/kg lead) ($p=0.009$). The percent change in body weight on Day 28 for the SlAb-A01 sample was also statistically different from the SlAb-A05 sample (260 mg/kg lead) ($p=0.027$).

The final histology report (presented in Appendix C) concluded that there was no toxicity associated with the field-collected (aged) soil exposures. No test article-related histopathologic findings were found.

Table 4–3 Summary of Soil Exposure Results – Survival and Growth

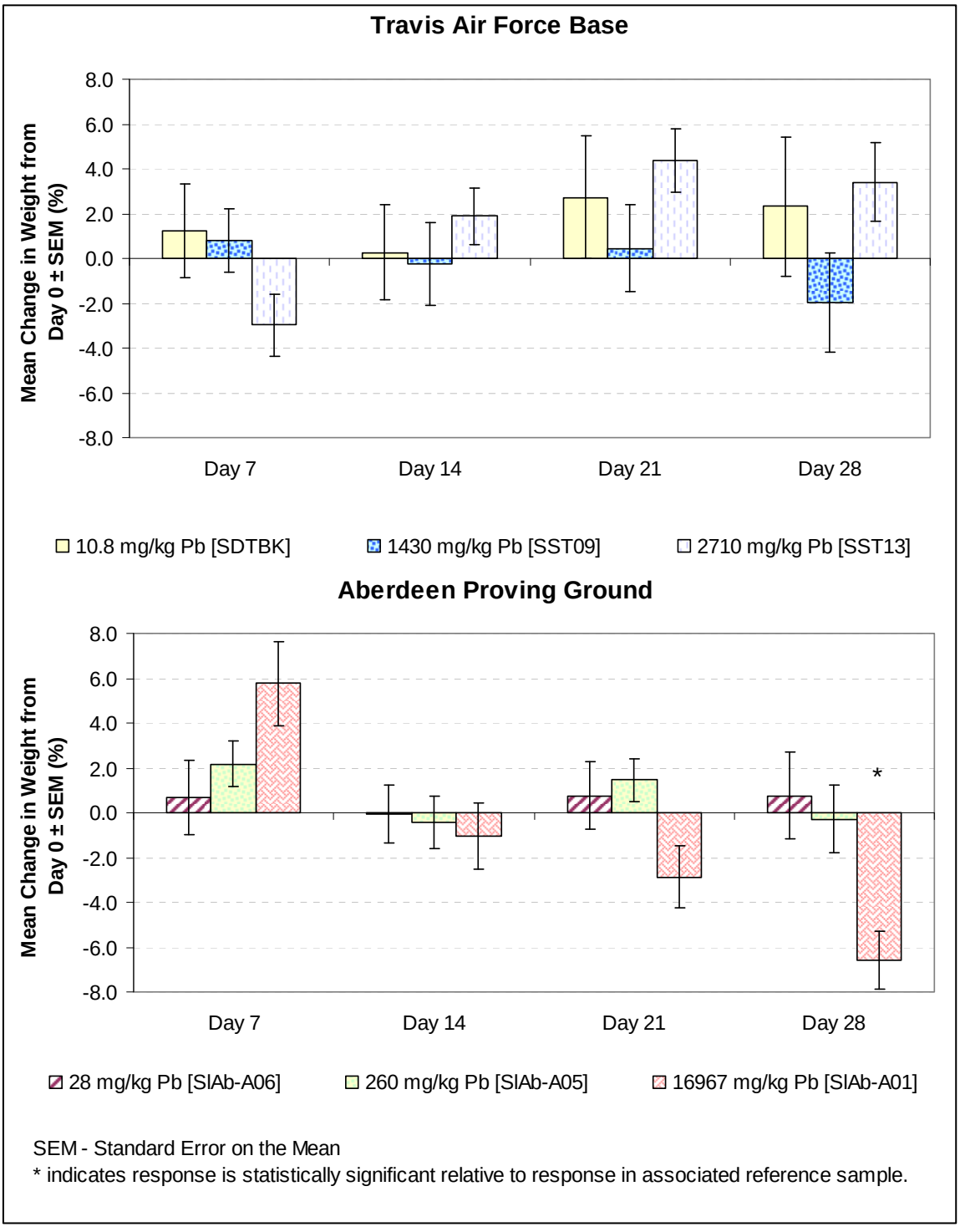
Location ID	Lead Concentration[1] (mg/kg)	TOC (%)	Mean Survival	Mean Body Weight (g)					Mean Change in Weight from Day 0 (%)			
			Day 28	Day 0	Day 7	Day 14	Day 21	Day 28	Day 7	Day 14	Day 21	Day 28
Travis Air Force Base												
SDTBK	10.8	1.5	100%	0.813	0.825	0.818	0.839	0.834	1.244	0.255	2.727	2.308
SST09	1430	1.6	100%	0.801	0.805	0.798	0.804	0.786	0.811	-0.250	0.461	-1.947
SST13	2710	2.3	100%	0.798	0.826	0.814	0.837	0.823	-2.975	1.882	4.360	3.412
Aberdeen Proving Ground												
SlAb-A06	28	1.3	100%	0.812	0.818	0.811	0.818	0.819	0.658	-0.073	0.748	0.755
SlAb-A05	260	0.41	100%	0.808	0.827	0.806	0.820	0.808	2.174	-0.447	1.461	-0.289
SlAb-A01	16967	0.088	100%	0.812	0.862	0.805	0.789	0.760	5.764	-1.025	-2.878	**-6.565**

1 - Samples analyzed by CHPPM on Days 0, 14, and 28 of test. Average is presented.
BK in location ID identifies background reference location; all others are site locations.
Bold text indicates result is statistically different from associated reference sample results.

Table 4–4 Summary of Soil Exposure Results – Blood Parameters

Location ID	Lead Concentration[1] (mg/kg)	TOC (%)	Average erythrocyte counts (10x4 cells/ul)	Average leukocyte counts (10x3 cells/ul)	Average Hemoglobin (g/dL)
Travis Air Force Base					
SDTBK	10.8	1.5	9.73	4.06	9.4
SST09	1430	1.6	9.40	4.20	8.8
SST13	2710	2.3	11.44	4.12	8.6
Aberdeen Proving Ground					
SlAb-A06	28	1.3	9.30	4.01	8.3
SlAb-A05	260	0.41	8.95	4.09	9.3
SlAb-A01	16967	0.088	8.98	4.80	8.4
1 - Samples analyzed by CHPPM on Days 0, 14, and 28 of test. Average is presented. BK in location ID identifies background reference location; all others are site locations. No results were statistically different from associated reference sample results.					

Figure 4-2 Mean % Change in Weight Over 28 Days of Soil Exposure



4.3.1.2 Evaluation of Field Demonstration Results Relative to Previous Studies

In general, the concentrations of lead in the field demonstration soils were similar to the range of concentrations evaluated in the laboratory validation testing. The maximum lead concentration from the APG demonstration site (16,967 mg/kg) was above the maximum concentration (9,167 mg/kg lead) detected in the lead-spiked soil evaluated in the laboratory validation phase of testing.

A comparison of the NOAECs and LOAECs derived from the laboratory validation tests conducted with lead acetate and the field demonstration tests show that less toxicity was observed in the test conducted with field-collected aged soils (Table 4-5). These results indicate that using screening values derived from studies conducted with laboratory-spiked soils may be overly protective of salamanders exposed to weathered metals under field conditions.

Table 4–5 Comparison of Soil NOAECs and LOAECs in Validation Testing and Field Demonstrations

Study	Lead (mg/kg)			
	Survival		Sub-Lethal Endpoint	
	NOAEC	LOAEC	NOAEC	LOAEC
Laboratory Validation [Spiked Soil]	1700	4700	1700	4700 [2]
Travis AFB	2710 [1]	>2710 [1]	2710 [1]	>2710 [1]
Aberdeen Proving Ground	16967 [1]	>16967 [1]	260	16967 [3]
All Field Demonstration Data	16967 [1]	>16967 [1]	2710	16967 [3]

Concentrations based on average of measured values at Days 0, 14, and 28 of test.

1 - No sub-lethal effects were observed at the highest tested concentration at demonstration site.

2 - Reduced leukocyte count observed at 4700 mg/kg lead.

3 - Reduced cumulative average growth observed at 16967 mg/kg lead.

For example, during the validation phase of testing, only 20% survival was recorded for salamanders exposed to the 9,167 mg/kg lead level, resulting in a survival NOAEC of 1,700 mg/kg lead. However, in the field demonstration testing, even at the highest lead concentration (16,967 mg/kg), no mortality was observed. This resulted in a survival NOAEC of 2,710 mg/kg lead for Travis AFB samples and 16,967 mg/kg lead for APG samples (the maximum tested concentration at each site). Since these NOAECs were derived based on a lack of observed toxicity, the actual NOAECs associated with a toxic response would be higher.

A similar trend was observed for the sub-lethal endpoints with toxicity observed at higher concentrations in the field demonstration tests. However, few statistically significant sub-lethal endpoints were observed in either the laboratory validation testing or the field demonstration testing. In the laboratory validation testing, the only sub-lethal response was a reduced leukocyte count observed in the 4,700 mg/kg lead treatment, resulting in a NOAEC of 1,700 mg/kg lead. The only statistically significant sub-lethal effect observed in the field demonstration was a reduction in

body weight (as percent change in body weight relative to Day 0) in the highest lead concentration (Figure 4-2).

Due to the wide distribution in lead concentrations for the APG site, the NOAEC for the percent change in body weight endpoint at APG was 260 mg/kg. It is likely that a higher NOAEC would be derived if additional samples between 260 mg/kg and 16,967 mg/kg lead were tested. If samples from both demonstration sites are evaluated together, the NOAEC is 2710 mg/kg lead (the highest tested sample from Travis AFB).

While these laboratory- and field-based NOAECs and LOAECs may be used as soil screening values, it is difficult to know the significance of the impact that some sub-lethal effects (e.g., reduced leukocyte counts) would have on the health of an individual salamander or a salamander population exposed to lead in the soil. A sub-lethal effect like reduced growth may be better related to population effects than a change in blood parameters.

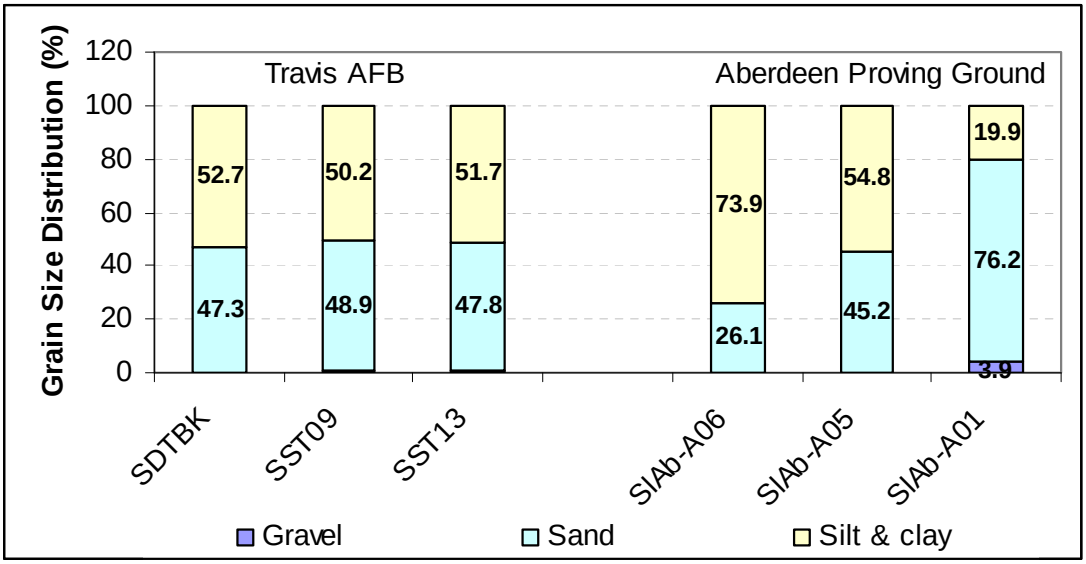
The difference in responses between the laboratory validation testing and the field demonstration testing may be explained by differences in the bioavailability of the lead. In the spiked-soil tests, the nature of the lead used to treat the soil (i.e., lead acetate) is such that the lead is likely to be highly bioavailable.

It appears that the weathering of the lead under field conditions may reduce the bioavailability of lead. Although the Travis AFB samples have weathered for less time than the APG samples (the skeet range at Travis AFB is still active), there did not appear to be a notable difference in bioavailability between the sites.

The level of TOC present in the soil may also have an effect on the observed toxicity with less toxicity expected in samples with higher TOC. The soil used in the laboratory validation testing contained 22.4% organic matter. Although TOC is a sub-set of the organic matter in a sample, there appears to be much more organic carbon in the laboratory validation soils than in the field demonstration soils (maximum TOC in tested soil samples was 2.3%).

The grain size composition may also be impact bioavailability. The laboratory validation soil was comprised of 45.6% sand, 43.6% silt, 10.8% clay. The composition of the laboratory validation soil is similar to the three Travis AFB samples and one of the APG samples (SlAb-A05) (Figure 4-3). The other APG samples had either more silt and clay (SlAb-A06) or more sand (SlAb-A01). It is not unexpected that the sandiest soil (SlAb-A01 at 76% sand) had the lowest TOC (0.088%), the highest lead level (16967 mg/kg) and the only observed toxic response.

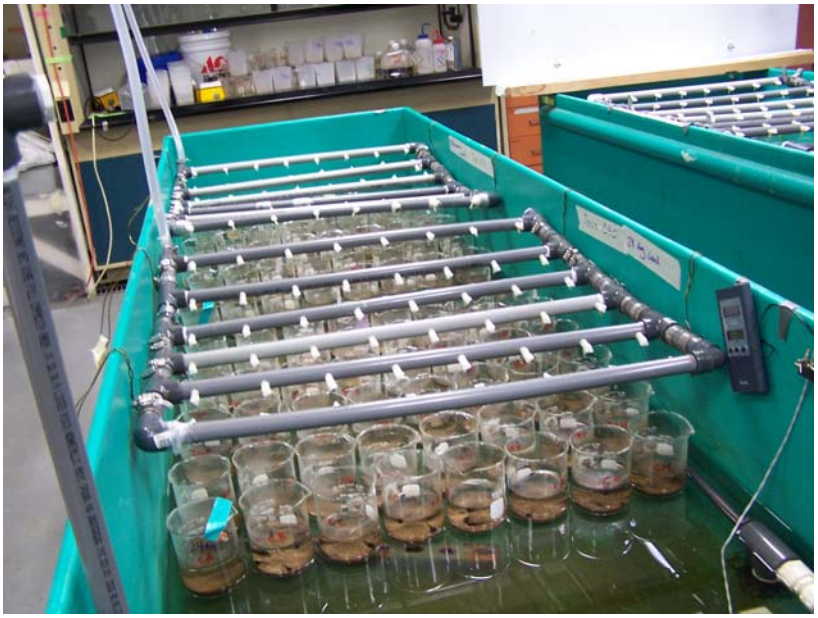
Figure 4-3 Grain Size Distribution of Tested Soils



4.3.2 Sediment Exposure Protocol

Tests were conducted for 10 days (in accordance with the protocol presented in Appendix A) with recently hatched tadpoles (Gosner Stages 17-20). Sediments were homogenized prior to placement in the test vessels. Each tested treatment consisted of eight replicates containing 10 larval tadpoles in each vessel. Figure 4-4 shows the set up of the toxicity test in the water bath (to maintain constant temperature) with the continuous drip flow-through system over the test vessels. Mortality and growth (i.e., body width, body length) were evaluated at test termination (Day 10).

Figure 4-4 Sediment Exposure Protocol Test Set-up



To select samples for the sediment toxicity test, the initial lead levels in each sample were reviewed (Table 3-11). In order to achieve a gradient of lead concentrations for the tests conducted for each demonstration site, dilutions of some sample were generated. For example, in the set of samples collected from the Travis AFB site, three dilutions of a sample with approximately 2000 mg/kg lead were made in order to achieve concentrations between approximately 80 and 850 mg/kg lead. One of the APG samples was also diluted to fill the gap between approximately 1000 and 17000 mg/kg lead.

Due to delays in test organism availability, the field demonstration tests were not started until December 2006 and January 2007. Copper and lead concentrations were analyzed in the fall of 2006 prior to the start of the tests. All sediments were stored in the dark at 4°C until testing began. A laboratory control was run with each set of tests to ensure the quality of the toxicity data.

4.3.2.1 Results of Field Demonstration

As described in Section 3.5.6, both lethal (mortality) and sub-lethal (growth as body width and body length) endpoints were measured in each toxicity study. Normality and homogeneity of variance were evaluated using the chi-square test and Bartlett's Test, respectively, with alpha = 0.01. Survival data were arcsine transformed, and growth data were not transformed. Statistical significance between the tested site samples and the reference site sample was evaluated using Steel's Many-One Rank Test (alpha = 0.05).

Treatments that demonstrated a significant reduction in survival (significantly higher mortality) were excluded from sub-lethal analyses since the sub-lethal measurements of surviving organisms can be skewed by a significant reduction in sample size. For the purposes of determining NOAECs and LOAECs these sub-lethal endpoints were treated as 'toxic' even though statistical analyses were not conducted. Table 4-6 presents the results of the sediment exposure tests for Travis AFB and APG. The results for each demonstration site are discussed below.

Eight samples (including diluted samples) from the Travis AFB site and a laboratory control were tested as part of the Travis AFB field demonstration. Lead concentrations ranged from 15 mg/kg in the reference sample to 3700 mg/kg. Copper concentrations were also analyzed in these samples to rule out another possible chemical stressor. Copper concentrations were low and not expected to be toxic. TOC concentrations were relatively consistent between samples, ranging from 1.5 % to 2.3%.

Survival in one sample from Travis AFB (SST07; 2100 mg/kg lead) was statistically lower than that observed for the reference sample (Figure 4-5). Since survival in this sample was significantly reduced, the sub-lethal endpoints were excluded from further statistical analyses. One sample (the M3 dilution of SST07) had a significant reduction in body length relative to the control.

Although toxicity was observed in the SST07 sample, samples with much higher lead levels did not show toxicity. The SDT14 and SST13 samples contained up to 3700 mg/kg lead with no impacts on survival or growth. Based on these results, the toxicity in the SST07 sample may not have been related to lead, or some characteristic of the sample increased the bioavailability of the lead present in the sample. Survival in the SST07 replicates was also variable with three replicates with total mortality (0% survival) and one replicate with no mortality (100% survival).

A review of the sample characteristics presented in Table 3-6 indicate that TOC, SEM and AVS, bulk density, CEC, and grain size values in the SST07 sample were similar to values for non-toxic samples. Levels of copper were also low in this sample. A review of the water quality parameters measured during the test (Appendix C) indicate that the dissolved oxygen (DO) in this samples was low (down to 3.4 mg/L), but within the acceptable range. Therefore, it is difficult to state with certainty that lead in the sediment was responsible for the toxicity observed in the SST07 sample. However, to be conservative, this sample was considered in the derivation of the LOAECs and NOAECs for the Travis AFB samples.

Eight samples (including diluted samples) from the APG site and a laboratory control were tested as part of the APG field demonstration. Lead concentrations ranged from 26 mg/kg in the reference sample to 17000 mg/kg. Copper concentrations were also analyzed in these samples to rule out another possible chemical stressor. Copper concentrations ranged from 7.7 mg/kg to 1200 mg/kg. During the laboratory validation phase of testing (summarized in Table 3-4), 87 mg/kg copper was sufficient to result in sub-lethal effects under low TOC conditions (TOC was 0.066%). Therefore, it is possible that, in addition to the lead, copper concentrations in the APG samples also contributed to observed toxicity. TOC levels in the APG samples were variable, ranging from 0.36% to 2.0%.

Survival in the SedAb-A3A sample and three dilutions of this sample was significantly reduced relative to the reference sample (Table 4-6; Figure 4-5). Since survival in these samples was significantly reduced, the sub-lethal endpoints were excluded from further statistical analyses. The SedAb-A04 sample had significant reductions in body width and length relative to the reference sample.

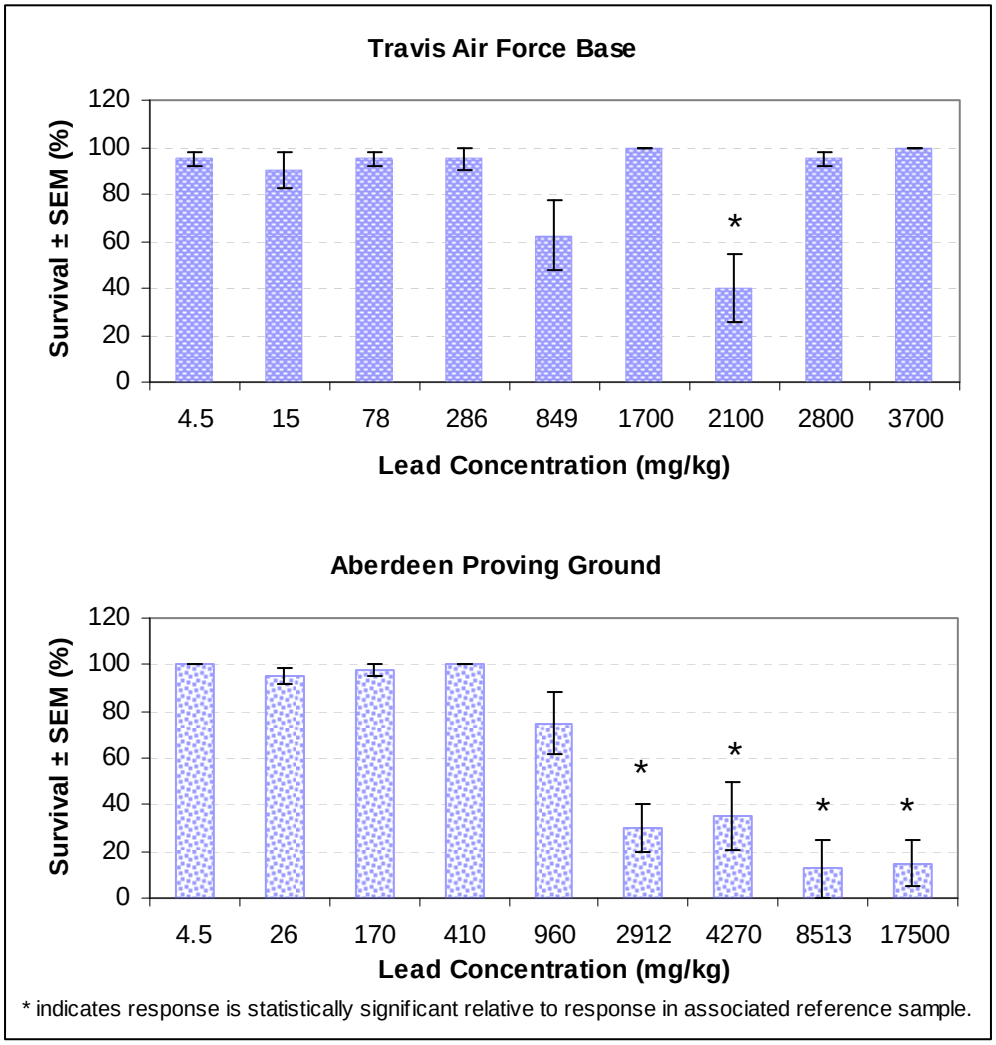
Table 4–6 Summary of Sediment Exposure Results

Location ID	Chemical Concentration [1] (mg/kg)		TOC (%)	Tadpole Results at Test Termination (Day 10)		
	Copper	Lead		Mean Survival (%)	Mean Body Width (mm)	Mean Body Length (mm)
Travis Air Force Base						
Lab Control	7.2	4.5	0.066	95	5.3	7.6
SDTBK	12	15	1.5	90	5.0	7.5
SST07 [M1]	12	78	1.5	95	5.1	7.5
SST07 [M2]	13	286	1.5	95	4.8	7.4
SST07 [M3]	14	849	1.5	62	3.4	**4.9** [2]
SDT04	19	1700	1.8	100	4.9	7.1
SST07	17	2100	1.6	**40** [2]	**2.5** [3]	**3.5** [3]
SDT14	21	2800	1.9	100	5.2	7.3
SST13	13	3700	2.3	95	5.0	7.0
Aberdeen Proving Ground						
Lab Control	7.2	4.5	0.066	100	5.4	9.0
SedAb-ABk11	7.7	26	0.46	95	5.6	9.7
SedAb-A08	16	170	0.36	98	5.7	9.8
SedAb-A07	37	410	1.9	100	5.8	9.9
SedAb-A04	140	960	0.57	75	**3.8** [2]	**6.0** [2]
SedAb-A3A [M1]	210	2912	0.72	**30** [2]	**1.3** [3]	**2.1** [3]
SedAb-A3A [M2]	306	4270	0.84	**35** [2]	**1.5** [3]	**2.4** [3]
SedAb-A3A [M3]	604	8513	1.2	**12** [2]	**0.4** [3]	**0.6** [3]
SedAb-A3A	1200	17000	2.0	**15** [2]	**0.4** [3]	**0.7** [3]

1 - Samples re-analyzed by analytical laboratory prior to toxicity testing .
2 - Indicates result is statistically different from reference sample results.
3 - Excluded from statistical analysis because survival was significantly reduced.
BK or Bk in location ID identifies background reference location; all others are site locations.

M1, M2, M3 concentrations achieved by diluting the following samples-
- for Travis Air Force Base - SST07 diluted with SDTBK
- for Aberdeen Proving Ground - SedAb-A3A diluted with SedAb-ABk11
- Copper, lead and TOC concentrations for these diluted samples are estimated based on the analytical results for the samples included in the dilution

Figure 4-5 Summary of Tadpole Survival Results



4.3.2.2 Evaluation of Field Demonstration Results Relative to Previous Studies

In general, the concentrations of lead detected in the field demonstration sediments were higher than the concentrations evaluated in the laboratory validation testing. The maximum tested lead concentrations from the Travis AFB demonstration site (3700 mg/kg) and the APG demonstration site (17500 mg/kg) were well above the maximum concentration (1200 mg/kg lead) evaluated in the lead-spiked sediments during the laboratory validation phase of testing.

Table 4-7 presents a comparison of the NOAECs and LOAECs derived from the laboratory validation tests and the field demonstration tests. These results show that less toxicity was observed in the field demonstration testing with aged sediment than with the spiked sediments in the validation testing.

Table 4–7 Comparison of Lead Sediment Screening Values in Validation Testing and Field Demonstrations

Study	Lead (mg/kg)			
	Survival		Sub-Lethal Endpoint [1]	
	NOAEC	LOAEC	NOAEC	LOAEC
Laboratory Validation [Spiked Sediment]	1200	>1200 [2]	100	260
Travis AFB	1700	2100 [3]	286	849 [3]
Aberdeen Proving Ground	960	2912	410	960
All Field Demonstration Data	1700	2100 [3]	410	849 [3]

Concentrations based on measured values prior to test initiation.

1 - Based on lower of values for body width and body length endpoints.

2 - No lethal effects were observed at the highest tested concentration.

3 - Acceptable survival and growth results were observed above this concentration; therefore these values may be overly conservative depending on site-specific conditions.

Since no mortality was observed in the maximum tested concentration during the laboratory validation phase of testing, the survival LOAEC was >1200 mg/kg lead. Lethal effects were observed in samples from both demonstration sites (Figure 4-5) resulting in survival LOAECs of 2100 mg/kg and 2912 mg/kg for the Travis AFB and APG sites, respectively. As indicated in Section 4.5.1, it is unclear whether lead is the stressor responsible for the observed toxicity in the Travis AFB sample since samples with higher lead levels did not show a significant reduction in survival (Figure 4-5). Although the survival NOAEC for the APG site is lower than that derived during the laboratory validation testing (1200 mg/kg lead), this is an artifact of the concentration gradient tested during the demonstration testing. No concentrations between 960 mg/kg and 2912 mg/kg were tested during the APG demonstration so the 960 mg/kg value became the NOAEC when survival was impacted in the 2912 mg/kg sample. If all of the sediment demonstration data are considered, the survival NOAEC becomes 1700 mg/kg lead.

The sub-lethal NOAECs and LOAECs were also higher in the field demonstration testing than the laboratory validation testing further indicating that lead is likely less bioavailable under field conditions (Table 4-7). As with the survival results, it is difficult to identify the stressors responsible for the sub-lethal impacts observed in the Travis AFB samples since samples with higher lead levels did not show a reduction in growth endpoints. Therefore, the sub-lethal LOAEC of 849 mg/kg based on the Travis AFB data may be overly conservative. There is more confidence in the sub-lethal LOAEC derived based on the APG results (960 mg/kg) since this data followed a more traditional dose-response curve (i.e., less growth observed in samples with higher lead levels).

These results indicate that using screening values derived from studies conducted with laboratory-spiked sediments is likely to over-estimate potential risks to amphibians. While this level of conservatism is appropriate for screening level risk analysis, the use of these screening values at wetland sites could result in overly conservative cleanup levels requiring remediation of larger wetland areas than may be warranted. The use of the sediment exposure protocol would provide a

site-specific assessment of the bioavailability and toxicity of lead, or other stressors, on larval amphibians that might be present in the wetland.

4.3.3 Assessment of Amphibian ERA Protocol

In 2004 NAVFAC published a guidance manual presenting the framework for a tiered amphibian ERA protocol that could be used to assess potential risks to amphibians as part of site evaluations at DoD (NAVFAC, 2004).

Conducting ERAs in a tiered, step-wise manner allows the risk assessor and risk manager to maximize the use of available site information and sampling data, while providing the opportunity to reduce the uncertainties inherent in the ERA process through the use of focused supplemental data collection to fill key data gaps identified in the previous tier of the assessment, if necessary. The Navy endorses a tiered approach for ERA (US Navy, 1999) and a tiered approach is consistent with USEPA methods for ERA (USEPA, 1997).

Although formal risk assessments were not completed for either of the field demonstration sites, the tiered approach may be applied using the available data to assess the versatility and applicability of the amphibian ERA protocol.

The Tier I amphibian ERA protocol essentially comprises a screening level ERA. This approach uses readily available information to identify potential amphibian exposure pathways at a site; determine which exposure pathways are complete; and conduct an effects-based screening using available benchmarks to determine whether or not the complete exposure pathways have the potential to pose a significant environmental risk.

The site reconnaissance efforts conducted at each of the sites provided an initial evaluation of amphibian habitat quality. The site selection effort focused on sites that had documented amphibian populations or appeared to contain amphibian habitat in order to assess sites with complete exposure pathways.

The initial habitat assessment at Travis AFB indicated that the site provides suitable habitat for at least one amphibian species. At the time of field sampling in March 2006, the average pool depth was approximately 50 cm with deeper areas over 80 cm. The surface area of the pool was approximately 1,200-square meters. The pool was primarily vegetated with annual hydrophytic grasses; curly dock (*Rumex sp.*), star thistle (*Centaurea solstitialis*), and other herbaceous materials. A Pacific tree frog (*Hyla regilla)* egg mass and several Pacific tree frog tadpoles were observed within the pool during a February 8, 2006 site visit. The calls of at least one adult Pacific tree frog were also heard.

During the March 2006 sampling effort, an ecological inventory was conducted using dip nets and meander surveys. No egg masses were observed and no amphibian species were collected using dip nets.

The initial habitat assessment at the Aberdeen Proving Ground study area indicated that the site provides suitable habitat for numerous amphibian species and other facultative vernal pool species.

At the time of sampling in April 2006, the average pool depth was approximately 60 to 90 cm. The pool was primarily vegetated with soft rush (*Juncus effusus*), broad-leaved cattail (*Typha angustifolia*), sedges (*Carex* spp.), and marsh bedstraw (*Galium palustre*). Ecological inventories at the site included dip netting, meander surveys, hand captures, and egg mass counts. The following species were observed at the site:

- Green frog (*R. clamitans*): vocalizations, hand capture
- Northern Cricket Frog (*Acris crepitans*): hand capture
- Spotted salamander (*Ambystoma maculatum*): 10 egg masses
- American toad (*B. americanus*): vocalizations
- Pickerel frog (*R. palustris*): vocalizations
- Fairy shrimp (*Eubranchipus* sp.): dip net
- Isopoda: dip net
- Unknown water beetle (*Coleoptera*): dip net
- Fishing spider (*Dolomedes* sp.): dip net

An effects-based screening was then conducted using the XRF data that was obtained as part of the reconnaissance effort at each site. Since copper was not detected in the Travis AFB XRF survey and copper XRF data were not collected the APG site, this initial screening focused on the available lead data. Table 4-8 compares the range of lead concentrations within each study area (excluding the reference locations) against literature-based soil and sediment screening values that might typically be used in an ERA conducted for a wetland site. These values were generally derived for the protection of terrestrial (e.g., earthworm, bird) and benthic (e.g., invertebrate) receptors, not for amphibians. The amphibian-based soil and sediment screening values from the laboratory validation phase of testing are also included in the table.

The maximum and average lead concentrations at each site were above the available soil and sediment screening values. The range of lead concentrations at each site was also above the XRF results for the associated reference locations. According to the Tier I amphibian ERA protocol, the presence of complete exposure pathways and concentrations above screening values and background locations would indicate that additional evaluation is warranted in the Tier II assessment.

The Tier II amphibian ERA protocol comprises a refined ERA using site-specific information to evaluate complete exposure pathways and amphibian ecological resources. This protocol provides quantitative measures and/or risk estimates of potential ecological effects associated with amphibian exposure to chemical stressors.

The soil and sediment sampling and analyses efforts conducted at each demonstration site after the XRF surveys would be consistent with the Tier II ERA process. Table 4-9 compares the analytical chemistry results for both copper and lead in field-collected samples (as opposed to the field-measured XRF data presented in Table 4-8) against soil and sediment screening levels indentified in the literature and derived during the validation phase of this project. Table 4-9 indicates that copper levels at the Travis AFB site are low and would not be expected to be toxic to worms, birds, salamanders, or tadpoles. However, levels of copper at APG and lead at both sites exceeded at least some of the screening values, indicating the potential for risk to wetland receptors.

Table 4–8 Screening of Field Demonstration Site XRF Survey Lead Data against Screening Values

				Lead (mg/kg)	
				Travis Air Force Base	**Aberdeen Proving Ground**
Field Demonstration XRF Survey Data			Minimum	464	38
			Average	1837	2265
			Maximum	4025	12387
			Background	20	33
Medium	**Source**	**Value**	**Receptor**	**Travis Air Force Base**	**Aberdeen Proving Ground**
Soil Screening Values	**From Literature**	Eco-SSL	Terrestrial invertebrate	**1700**	**1700**
		Eco-SSL	Vertebrate [bird]	**11**	**11**
	From Validation Testing	Survival NOAEC	Salamander	**1700**	**1700**
		Sub-Lethal NOAEC	Salamander	**1700**	**1700**
Sediment Screening Values	**From Literature**	TEC	Benthic invertebrate	**35.8**	**35.8**
	From Validation Testing	Survival NOAEC	Tadpole	**1200**	**1200**
		Sub-Lethal NOAEC	Tadpole	**100**	**100**

Boldface indicates that maximum site concentration exceeds that screening value.
Validation testing values were presented in Tables 3-3 and 3-4 and are based on measured concentrations.
Copper was either not analyzed (APG) or not detected (Travis APG) in the XRF surveys.
XRF data are not corrected for percent moisture (corrected values would be approximately 25% lower; assuming percent solid is approximately 75%).
Eco-SSL - Ecological-Soil Screening Level (USEPA, 2005b; USEPA, 2007). Vertebrate Eco-SSL is the lower of the avian and mammalian Eco-SSLs.
NOAEC - No Observed Adverse Effect Concentration
TEC - Threshold Effect Concentration (MacDonald, et al, 2000)

Table 4–9 Screening of Field `Demonstration Site Analytical Data against Screening Values

				Travis Air Force Base		Aberdeen Proving Ground	
				Copper (mg/kg)	**Lead (mg/kg)**	**Copper (mg/kg)**	**Lead (mg/kg)**
Field Demonstration Analytical Data			Minimum	16	935	17	35
			Average	18	2117	232	3166
			Maximum	21	4200	1150	17500
Medium	**Source**	**Value**	**Receptor**	**Copper (mg/kg)**	**Lead (mg/kg)**	**Copper (mg/kg)**	**Lead (mg/kg)**
Soil Screening Values	**From Literature**	Eco-SSL	Terrestrial invertebrate	80	**1700**	**80**	**1700**
		Eco-SSL	Vertebrate [bird]	28	**11**	**28**	**11**
	From Validation Testing	Survival NOAEC	Salamander	803	**1700**	**803**	**1700**
		Sub-Lethal NOAEC	Salamander	2700	**1700**	2700	**1700**
Sediment Screening Values	**From Literature**	TEC	Benthic invertebrate	31.6	**35.8**	**31.6**	**35.8**
	From Validation Testing	Survival NOAEC	Tadpole	230	**1200**	**230**	**1200**
		Sub-Lethal NOAEC	Tadpole	21	**100**	**21**	**100**

Boldface indicates that maximum site concentration exceeds that screening value.
Validation testing values were presented in Tables 3-3 and 3-4 and are based on measured concentrations.
Eco-SSL - Ecological-Soil Screening Level (USEPA, 2005b; USEPA, 2007). Vertebrate Eco-SSL is the lower of the avian and mammalian Eco-SSLs.
NOAEC - No Observed Adverse Effect Concentration.
TEC - Threshold Effect Concentration (MacDonald, et al, 2000).

At this point in the Tier II ERA, depending upon stakeholder concerns and regulatory status, additional work might be recommended to further assess potential impacts to amphibians. This could be accomplished using a variety of methods including laboratory toxicity tests or field surveys designed to assess the diversity and abundance of the existing amphibian community. In some cases, bioaccumulation evaluations (e.g., site-specific tissue collection, laboratory exposures) may be warranted if risks to higher trophic level receptors are of concern.

For the field demonstration, laboratory toxicity tests were used to directly assess the bioavailability and toxicity of the soil and sediment samples. As described in Sections 4.3.1.1 and 4.3.2.1, the sub-lethal effects on salamanders and both lethal and sub-lethal impacts on larval amphibians were observed during the testing. In the Tier II ERA, these responses could be used to identify toxic sampling locations that might need further study or remediation or, since a lead concentration gradient was tested, to establish clean up levels based on the NOAECs and LOAECs.

Table 4-10 compares the results of the field demonstration test against the soil and sediment screening values identified in the literature and the screening values derived using spiked soil and sediment in the laboratory validation phase of testing. The field demonstrations observed less toxicity than would have been predicted using screening values typically used in ERAs (e.g. literature based screening values for plants, terrestrial invertebrates, or benthic invertebrates) or using amphibian toxicity data generated using spiked soils and sediments.

This demonstration indicates that the application of the tiered amphibian ERA protocol and laboratory testing with the soil and sediment exposure protocols was appropriate for assessing potential risks to amphibians at both demonstration sites. At both demonstration sites, the use of technologies developed and refined through the ESTPC program identified less potential for risk to amphibians, and therefore less area potentially requiring remediation, than would have been identified by applying the literature based screening levels that have previously been used in wetlands.

Table 4–10 Comparison of Demonstration Testing Results and Screening Values for Lead

Medium	Source	Value	Receptor	Lead (mg/kg)
Soil Screening Values	**From Literature**	Eco-SSL	Terrestrial invertebrate	1700
		Eco-SSL	Vertebrate [bird]	11
	From Validation Testing	Survival NOAEC	Salamander	1700
		Sub-Lethal NOAEC	Salamander	1700
	From Demonstration Testing	Survival NOAEC	Salamander	2710 (TAFB) 16967 (APG)
		Sub-Lethal NOAEC	Salamander	2710 (TAFB) 260 (APG)
Sediment Screening Values	**From Literature**	TEC	Benthic invertebrate	35.8
	From Validation Testing	Survival NOAEC	Tadpole	1200
		Sub-Lethal NOAEC	Tadpole	100
	From Demonstration Testing	Survival NOAEC	Tadpole	1700 (TAFB) 960 (APG)
		Sub-Lethal NOAEC	Tadpole	286 (TAFB) 410 (APG)

Validation testing values were presented in Tables 3-3 and 3-4 and are based on measured concentrations.
Demonstration testing values were presented in Tables 4-5 and 4-7 and are based on measured concentrations.
TAFB - Travis Air Force Base demonstration testing results.
APG - Aberdeen Proving Ground demonstration testing results.
Eco-SSL - Ecological-Soil Screening Level (USEPA, 2005b; USEPA, 2007). Vertebrate Eco-SSL is the lower of the avian and mammalian Eco-SSLs.
NOAEC - No Observed Adverse Effect Concentration.
TEC - Threshold Effect Concentration (MacDonald, et al, 2000).

5.0 Cost Assessment

In addition to assessing the technical performance of the amphibian testing protocols, developing an understanding of cost performance is equally important. Cost considerations to be reported and evaluated include the perceived "real" costs associated with implementing the amphibian testing protocol as part of a larger site characterization effort. These costs are readily quantifiable and are based on site-specific conditions, including but not limited to the regulatory status of the site, size of the impacted site, number of samples, and laboratory testing requirements. Section 5.1 presents additional detail regarding this element of the cost assessment.

In addition, to "real" costs, use of technologies such as the amphibian testing protocol also has "opportunity" cost implications. When the toxicity testing protocols are appropriately applied, the user may avoid potential opportunity cost(s) associated with using a more conservative risk management approach. For instance, the use of inappropriate site characterization technologies in a palustrine wetland may result in costly and unnecessary wetland remediation based on the use of inappropriate endpoints. A more detailed discussion related to the opportunity cost savings presented by this innovative technology are presented in Section 5.2.

5.1 Cost Analysis

A summary of the approximate range of costs associated with implementing the amphibian testing protocols at several progressively sized sites is provided below in Table 5-1. This table further quantifies the use of a tiered amphibian ERA approach presented in a guidance manual published by NAVFAC (NAVFAC, 2004). Tier I of the amphibian ERA protocol represents a screening level ERA, which uses readily available information to identify potential amphibian exposure pathways. The results of the Tier I screening level ERA are typically used to determine whether or not additional amphibian ERA is warranted. Should the results of the Tier I assessment indicate that further amphibian ERA activities are not warranted, the Tier I activities would represent a finite and typically a *de minimus* costs for the end user, in relation to the overall site characterization. In this scenario, the costs associated with the Tier I screening level ERA would represent the extent of costs associated with the application of the amphibian testing technology at a site.

The Tier II portion of the protocol is a refined ERA, and is conducted to evaluate site-specific exposure pathways recommended at the conclusion of the Tier I evaluation. The need for additional sampling to evaluate potential risks to amphibians must be reviewed in terms of project-specific objectives. Additional data needs may include sampling and analysis of additional sediment, hydric soil, or surface water samples from within the study area or appropriate background locations. Depending upon site-specific circumstances, collection of sediment or hydric soil for laboratory toxicity testing may also be required. In addition, site-specific amphibian field studies may be warranted. These studies may include determining what amphibian species occur at the site, the relative abundance of those species, and collecting and analyzing amphibian tissue. Amphibian field survey results may be compared relative to reference sites to determine if measured concentrations of chemicals in abiotic media are related or correlated with field observations.

Table 5–1 Tier I and Tier II Amphibian ERA Implementation Costs

Cost Category	Sub Category	Details	Costs		
			Site A	Site B	Site C
Tier I ERA Costs					
Screening Costs	Site Characterization/ Screening Level ERA	Review of available information	$ 7,500	$ 17,500	$ 37,500
Tier II ERA Costs					
Start-up Costs	Site Reconnaissance	Labor and travel for 2 people	$ 1,790	$ 3,880	$ 5,970
	Mobilization	Planning, contracting, site preparation, personnel mobilization, supply shipping	$ 4,000	$ 5,000	$ 5,500
Capital Costs	Capital Equipment Purchases	Sampling/ homogenizing	$ 400	$ 900	$ 1,800
Direct Operating Costs	Capital Equipment Rentals	XRF analyzer	$ 600	$ 1,200	$ 3,000
	Toxicity Testing	Amphibian toxicity ($1200/sample)	$ 4,800	$ 10,800	$ 21,600
	Supervision	Labor and travel for 1 person	$ 500	$ 1,145	$ 3,080
	Operator Labor	Labor and travel for 2 people	$ 1,790	$ 3,880	$ 10,150
	Consumables/ Supplies	Sampling/ decontamination	$ 800	$ 1,620	$ 3,240
	Sampling and Analysis	Chemistry analyses ($425/sample)	$ 3,400	$ 7,650	$ 15,300

Table 5–1 Tier I and Tier II Amphibian ERA Implementation Costs (continued)

Cost Category	Sub Category	Details	Costs		
			Site A	Site B	Site C
Indirect Operating Costs	Environmental and Safety Training	OSHA 40 hour training for 2 samplers ($600/person)	$ 1,200	$ 1,200	$ 1,200
Demobilization	Demobilization	Equipment decontamination, shipment of supplies, personnel demobilization	$ 2,000	$ 2,500	$ 2,750
Other	Report Preparation	Evaluate potential for risk and establish remedial goals	$ 20,000	$ 25,000	$ 30,000
Total Implementation Costs					
Cost of Tier I and Tier II ERA			$ 48,780	$ 82,275	$ 141,090

Site A = 2 acres; 4 toxicity testing samples; 8 analytical samples; 1 day of site reconnaissance; 1 day of field sampling
Site B = 15 acres; 9 toxicity testing samples; 18 analytical samples;2 days of site reconnaissance; 2 days of field sampling
Site C = 30 acres; 18 toxicity testing samples; 36 analytical samples; 3 days of site reconnaissance; 5 days of field sampling
All costs are estimates and could vary by up to 50% depending upon site-specific conditions.
Chemical analyses include metals, TOC, grain size, and SEM/AVS.

Assumptions:

8 hour field days with 2 field staff	Travel assumptions:
Field staff rate = $100/hour	No airfare included
Supervisor rate = $150/hour	Hotel = $150/night
Supervisor in the field 50% of the time	Car + mileage = $90/day
XRF rental fee is $600/day	Meals = $50/day

When the early life stage frog (sediment) bioassay protocol is used at a site, as with other toxicity testing procedures, the unit costs are expected to vary somewhat based on market conditions, number of tests being considered, nature of contamination, and other site-specific considerations. The expected costs to implement the 10-day amphibian toxicity testing protocol (ASTM E2591-07 Standard Guide for Conducting Whole Sediment Toxicity Tests with Amphibians) generated through this ESTCP program are expected to be similar to other ASTM and USEPA assays such as the 10-day benthic invertebrate toxicity tests conducted with the midge, *Chironomus tentans,* and the amphipod, *Hyalella azteca*. Actual unit costs for these benthic invertebrate assays (in 2008 dollars) range from approximately $750 to $1500 per 10-day test, depending upon site-specific circumstances, whereas longer term tests are typically proportionately scaled. It is anticipated that the amphibian testing protocol market costs will be within ± 20% of the invertebrate costs.

The costs to implement the amphibian ERA protocol is primarily dependent upon the spatial scale of the area under investigation and the number of samples required to meet the data quality objectives. For the sediment exposure protocol, the duration of the toxicity test can be increased to allow the evaluation of additional sub-lethal endpoints, and this increase in duration will have an impact on the implementation costs. Once the spatial scale of the area has been established, cost drivers are expected to be primarily related to labor, travel, laboratory analytical costs, and laboratory toxicity testing costs, which will vary from site to site.

The size of the site under investigation provide a basis for the number of personnel hours required to conduct the field surveys and collect the soil and/or sediment samples for evaluation. The number of samples submitted for analytical or toxicological evaluation will likely increase with the size of the site and will impact the amount of labor needed to conduct the analyses and the toxicity tests, as well as the level of effort associated with the evaluation of the associated results and generation of the project reports.

The distance of the site from airports, hotels, and the field team's home base will increase costs if the area under investigation is relatively isolated or distant. Costs associated with mobilizing and demobilizing equipment for the field effort are largely dependent upon labor and shipping costs. Labor is likely to be relatively consistent from site to site, but shipping costs, like travel, will vary depending upon distance to the site and method of transportation.

As the size of the site increases, the per sample incremental costs associated with travel, reporting, mobilization and sample collection are driven down by efficiencies associated with economies of scale. For example, Table 5-1 provides a range of costs to conduct the amphibian ERA at three sites with varying acreage and equivalent conditions as they relate to costs (i.e., location from field team base, analytical parameters, labor rates). The savings associated with a larger site can be viewed on a unit basis by dividing the total cost per site by the acreage or samples to be collected and presenting the costs on a per acre or per sample basis, as presented in Table 5-2.

Table 5–2 Incremental Implementation Costs

	Amphibian ERA Incremental Costs	
	Per acre	**Per toxicity testing sample**
Site A[1]	$ 24,390	$ 12,195
Site B[1]	$ 5,485	$ 9,142
Site C[1]	$ 4,703	$ 7,838
1 – Total costs for conducting Tier I and Tier II ERA at each site are detailed in Table 5-1.		

5.2 Opportunity Cost Evaluation

As previously discussed, the cost implications associated with implementing the amphibian ERA protocol as a means to derive ERA-based remedial goals are two dimensional. In many cases, alternative, non-wetland ecological receptors are inappropriately used to derive ERA-based remedial goals at wetland sites. The use of these organisms has the potential to overestimate potential risks and increase project costs, or alternatively to under-estimate potential risks, and thereby result in a less costly, but less protective, risk management decision.

In the absence of the amphibian sediment testing protocol, remedial risk-management decisions in wetlands often rely on site-specific benthic invertebrate toxicity testing using organisms such as the amphipod, *Hyalella azteca,* or the midge, *Chironomus tentans*. While these species may not be present in many of the wetlands in questions, they are commonly accepted surrogates for assessing toxicity. Implementing the amphibian sediment testing protocol could be as much as 20% more costly than these traditional methods (depending upon site-specific circumstances). However, the value in expending this additional amount is achieved when making an informed decision about incurring the financial burdens associated with unnecessary wetland remediation and the preventable loss of valuable wetland resources.

The costs associated with using an inappropriate ERA-based remedial goal to require unnecessary environmental activities has four major cost implications, including: the derivation and negotiation of clean-up goals, the remediation activities, the wetland restoration activities, and the more intangible disturbance associated with disturbing the wetland.

The DoD has historically expended considerable effort and time attempting to assess impacts to amphibians or negotiating more reasonable remedial goals than the ecological screening levels that could serve as an initial overly conservative remedial goal. At Site 22, a 500 acre munitions bunker area in the Inland Area of Naval Weapons Station Seal Beach Detachment Concord in Concord, California, the endangered California tiger salamander (*Ambystoma californiense*) has been identified as an ecological receptor with the potential for exposure to arsenic in shallow soil. However, because there is not an ecological screening value for salamanders exposed to arsenic in soil it has been difficult to quantitatively evaluate the risk to these receptors. The project schedule and budget have been impacted by requests from the regulatory agencies to quantitatively assess risks to the salamanders in the absence of an appropriate soil screening value or an accepted

methodology. This issue has led to an extended comment resolution process on documents, and the project team has expended considerable effort to resolve these comments and stay within the Federal Facilities Agreement (FFA) schedule at this National Priorities List (NPL) site. The risk assessment challenge at this site exemplifies the need for amphibian-based ecological risk assessment methodologies and testing protocols for soil.

The sediment toxicity test protocol using northern leopard frog tadpoles (*R. pipiens*) was included, along with midge sediment toxicity tests, in the 2005 Baseline Ecological Risk Assessment conducted for Tributary 2 of Operable Unit 1 (OU-1) at Cherry Point, North Carolina (CH2M Hill, 2005). Contaminants measured in the sediments included heavy metals, PAHs, pesticides, VOCs, and SVOCs. The sediment toxicity test offered a means to directly evaluate potential risks to amphibians, instead of using other organisms (i.e., aquatic or sediment invertebrates) as surrogates. The results of the toxicity tests indicated that potential impacts to amphibians were expected to be minimal and that potential risks to the midge were greater. At this site the amphibian data were used to show that amphibians were not an at-risk receptor group and that risk management efforts and remediation should focus on the benthic macroinvertebrate community. The use of the amphibian test results was considered "cost effective uncertainty reduction" since it gave the project team site-specific amphibian data on which conclusions could be drawn.

At the Naval Weapons Station (NWS) Yorktown site in York County, Virginia, the sediment toxicity test protocol was included in a toxicity testing program designed to generate preliminary remediation goals for metals (e.g., mercury, arsenic, cadmium, and selenium, silver) found in a palustrine scrub/shrub wetland. The toxicity testing program included testing with green frog tadpoles (*R. clamitans*), the amphipod (*Hyalella azteca*), and the fathead minnow (*Pimephales promelas*). Although remediation has not yet occurred, the arsenic NOAEC from the amphibian test and the mercury NOAEC from the amphipod test will likely be used to help determine the remedial action.

Remediation costs can and will vary significantly from one site to another. Factors such as the type of contaminants, contaminant concentrations, the three dimensional nature of impacts in the subsurface, leachability of the contaminants, accessibility of the site, and local resources available to perform remedial activities can all play a major role in the total remediation costs. Due to the wide variety of factors that can affect remediation costs it is impossible to provide a narrow range since costs can easily range from several thousand to millions of dollars.

Wetland restoration costs vary regionally and by complexity and wetland type. The most costly restoration efforts involve significant soil management activities (i.e., excavation, disposal, backfill, and grading) and hydrologic manipulation (i.e., dewatering, water treatment and disposal, stream diversion, extraction wells, etc.). Wetland restoration costs involving only limited backfill and grading to replace an herbaceous emergent wetland can range from $40,000 to $80,000/acre (reflecting regional variation), while the costs for restoration of a palustrine scrub-shrub or forested wetland complex requiring 2 feet of backfill and hydrologic modifications during construction may approach $85,000 to $135,000/acre. If riparian corridor/stream restoration and the associated armoring or bioengineering structures are also required, costs (excluding soil management and disposal) can range up to $150,000/acre. In comparison, applying the amphibian risk assessment at a 10-acre forested palustrine wetland site would cost approximately $20,000 to $100,000 (depending upon site-specific considerations), and potentially result in a no action finding based on

use of technically appropriate risk assessment endpoints. In comparison, the potential ecological restoration costs (not including soil or sediment management costs, which might even outweigh restoration costs) in the same wetland system may be as high as $1.5 million.

Assigning a monetary value to the disturbance of an ecosystem/wetland when those activities are unwarranted is very difficult to quantify, yet the costs are real. Among the many valuable, but relatively intangible, benefits of a wetland ecosystem system include the improvement of water quality, flood control, recreation, shoreline erosion control, and a habitat for a multitude of species. When reviewing the costs of remediating or restoring a wetland, the ecological costs associated with the disturbance of the wetland habitat need to be considered.

6.0 Implementation Issues

6.1 Environmental Checklist

In general, under CERCLA status, the collection of soil or sediment from most locations would not require any local or state environmental permits. If threatened or endangered species are known to occur within a sampling site, additional care must be taken to avoid injuring protected species or their habitats. Under most state programs, a scientific collection permit would be required if a field program anticipated collection of amphibians for tissue analyses.

All participants in the field effort would be expected to comply with health and safety regulations and all facility-specific requirements while working at the sites.

6.2 Other Regulatory Issues

Neither the skeet shooting range at Travis AFB nor the small arms range at APG is currently part of site characterization or investigation activities. Both sites have been included in previous on-site investigations (i.e., a remedial investigation at Travis AFB, habitat surveys and screening level ERA at APG). The results of the demonstration could be incorporated into the evaluation of corrective actions if such actions are suggested by other site investigations.

6.3 End-User Issues

The primary end-users for the toxicity testing protocols will be site investigators and the regulators that review ecological risk assessments. Other stake-holders involved in the ERA process may include groups like the USFWS and the general public. The sediment exposure protocol has been included and accepted as a component of ERAs conducted at several locations including the Massachusetts Military Reservation and the soil exposure protocol has been used to develop soil screening values for several compounds of interest to the Army.

The sediment exposure protocol and the amphibian ERA framework are both applicable to investigating potential impacts to amphibians due to exposure to a variety of contaminants including metals, pesticides, PAHs, and PCBs associated with sediments or hydric soils in wetlands or other aquatic habitats that may occur on DoD facilities. The soil exposure protocol is expected to be most appropriate for controlled toxicological investigations designed to derive safe soil levels for particular compounds.

The equipment required for this technology (e.g., field survey equipment and laboratory supplies) will generally be commercially-off-the-shelf (COTS) items. However, not all environmental laboratories are set-up to run these types of assays or have experience with these test organisms. It is recommended that end-users thoroughly investigate the qualifications of the toxicity testing laboratory prior to conducting the soil and sediment protocols. The chemical analyses used to characterize the soil and sediment samples are typical for most site investigations. It is recommended that end-users identify a chemistry lab that is accustomed to analyzing samples from hazardous waste sites.

Technology transfer efforts have been on-going over the course of this project. The development of an approved ASTM sediment testing guide is one key component of the technology transfer and will be important in gaining acceptance of this technology by both regulatory agencies and end-users. As discussed in Section 1.2, the draft guide was submitted to the full Biological Effects and Environmental Fate committee in August 2007 and was approved in November 2007. The protocol was published in December 2007 and is presented in Appendix D. When a procedure has been accepted as an ASTM standard for conducting physical, chemical, or biological measurements, it inspires confidence among end users, and facilitates regulatory (e.g., USEPA) acceptance of innovative technologies.

The results of the validation phase of testing have been presented at several conferences and it is anticipated that the results of the field demonstration and species sensitivity testing will be presented at upcoming conferences. These conferences represent opportunities to present the results of this project and discuss the use of the amphibian protocol with site investigators and regulators. Several of these scientific conferences are attended by representatives from universities, federal and state government agencies, and environmental consulting firms from around the world and presenting the ESTCP project in these venues is an important part of publicizing the work and achieving regulatory acceptance.

Through these efforts and others the sediment toxicity testing protocol has been implemented at several DoD facilities (i.e., Cherry Point, North Carolina, Massachusetts Military Reservation, Massachusetts, and NWS Yorktown, Virginia) and private sites (i.e., Massachusetts Highway Department site and site investigation led by USEPA Region 4) under several different regulatory programs. Posters and presentations have been presented at the following venues:

- Tri-Service Ecological Risk Assessment Work Group (TSERAWG) Meetings in May 2005 and May 2006
- ESTCP/SERDP Symposia in December 2006 and 2007
- Society of Environmental Toxicology and Chemistry (SETAC), North America Annual Meeting in November 2006 and November 2008
- University of Massachusetts Annual Conference on Soils, Sediments and Water in October 2006
- In Situ and On-Site Bioremediation Symposium in May 2007
- DoD Operational Range Assessment and Management Meeting in August 2007.

Team members have also presented project information at an EPA Region 3 Biological Technical Assistance Group (BTAG) meeting, at a USGS Patuxent Wildlife Research Center (PWRC) seminar, in the AFCEE Technology Transfer Newsletter that is distributed to over 75,000 regulators, consultants, and members of the DoD, and in an upcoming issue of the Navy's magazine Currents. Presentations are also anticipated at Battelle's February 2009 Fifth International Conference on Remediation of Contaminated Sediments.

An article discussing the toxicological responses of red-backed salamanders (*P.cinereus*) to soil exposures of copper has been accepted by a peer-reviewed journal (Bazar, et al., 2008) and an article discussing the response of the salamanders to lead exposures is in progress.

7.0 References

Bazar, M.A., M.J. Quinn Jr, K. Mozzachio, J.A. Bleiler, C.R. Archer, C.T. Phillips, and M.S. Johnson. 2008. Toxicological responses of red-backed salamanders (*Plethodon cinereus*) to soil exposures of copper. (accepted September 2008). Archives of Environmental Contamination and Toxicology.

Chaney RL, Mielke HW, and Sterrett SB. 1989. Speciation, mobility and bioavailability of soil lead. In: BF. Davies, BG. Wixson editors. Lead in soil: Issues and guidelines. Environmental Geochemistry and Health, Monograph Series 4, Supplement to Volume 9. Northwood, UK: Science Reviews. p 105–129.

CH2M Hill, 2005. Baseline Ecological Risk Assessment Operable Unit 1 (OU-1). Marine Corps Air Station, Cherry Point, North Carolina. Prepared for Department of the Navy NAVFAC Atlantic. August.

Johnson, M.S., H.I. Paulus, C.J. Salice, R.T. Checkai, and M. Simini. 2004. Toxicologic and histopathologic response of the terrestrial salamander: *Plethodon cinereus* to soil exposures of 1,3,5-trinitrohexahydro-1,3,5-triazine (RDX). Archives of Environmental Contamination and Toxicology. 47:496-501.

Loehr, R.C., 1996. The Environmental Impact of Soil Contamination: Bioavailability, Risk Assessment, and Policy Implications. Reason Foundation. Los Angeles, California. Policy Study No. 211.

MacDonald, D.D., Ingersoll, C.G. and T.A. Berger. 2000. Development and Evaluation of Consensus-Based Sediment Quality Guidelines for Freshwater Ecosystems. Archives of Environmental Contamination and Toxicology. 39:20-31.

Naval Facilities Engineering Command (NAVFAC). 2004. Development of a Standardized Approach for Assessing Potential Risks to Amphibians Exposed to Sediments and Hydric Soils. March.

Naval Facilities Engineering Command (NAVFAC). 2005. Laboratory Validation Plan for ESTCP Project CU-0514 Demonstration and Certification of Amphibian Ecological Risk Assessment Protocol. December.

Naval Facilities Engineering Command (NAVFAC). 2006. Site Selection Memorandum for ESTCP Project CU-0514 Demonstration and Certification of Amphibian Ecological Risk Assessment Protocol. February.

Naval Facilities Engineering Command (NAVFAC). 2007a. Field Demonstration Plan for ESTCP Project CU-0514 Demonstration and Certification of Amphibian Ecological Risk Assessment Protocol. May.

Naval Facilities Engineering Command (NAVFAC). 2007b. Test Refinement Interim Report for ESTCP Project CU-0514 Demonstration and Certification of Amphibian Ecological Risk Assessment Protocol. June.

Thorbjornsen, K. and J. Myers. 2007. Identification of Metals Contamination in Firing-Range Soil Using Geochemical Correlation Evaluation. Soil and Sediment Contamination, Volume 16. 4:337-349.

US Environmental Protection Agency (USEPA). 1997. Ecological Risk Assessment Guidance for Superfund, Process for Designing and Conducting Ecological Risk Assessments (Interim Final). EPA 540/R-97/006. Office of Solid Waste and Emergency Response, Office of Emergency and Remedial Response. June.

US Environmental Protection Agency (USEPA). 2005a. Procedures for the Derivation of Equilibrium Partitioning Sediment Benchmarks (ESBs) for the Protection of Benthic Organisms: Metal Mixtures (Cadmium, Copper, Lead, Nickel, Silver and Zinc). EPA-600-R-02-011. Office of Research and Development. January.

US Environmental Protection Agency (USEPA). 2005b. Ecological Soil Screening Levels for Lead: Interim Final. OSWER Directive 9285.7-70. Office of Solid Waste and Emergency Response, March.

US Environmental Protection Agency (USEPA). 2007. Ecological Soil Screening Levels for Copper: Interim Final. OSWER Directive 9285.7-68. Office of Solid Waste and Emergency Response, Issued July 2006, Revised February 2007.

US Navy. 1999. Navy Policy for Conducting Ecological Risk Assessments. Issued by the Office of the Chief of Naval Operations on April 5, 1999. Available at http://web.ead.anl.gov/ecorisk/policy/pdf/ecoriskpolicy.pdf

Appendix D

ASTM Standard Guide

Designation: E 2591 – 07

Standard Guide for Conducting Whole Sediment Toxicity Tests with Amphibians[1]

This standard is issued under the fixed designation E 2591; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This standard covers procedures for obtaining laboratory data concerning the toxicity of test material (for example, sediment or hydric soil (that is, a soil that is saturated, flooded, or ponded long enough during the growing season to develop anaerobic (oxygen-lacking) conditions that favor the growth and regeneration of hydrophytic vegetation)) to amphibians. This test procedure uses larvae of the northern leopard frog (*Rana pipiens*). Other anuran species (for example, the green frog (*Rana clamitans*), the wood frog (*Rana sylvatica*), the American toad (*Bufo americanus*)) may be used if sufficient data on handling, feeding, and sensitivity are available. Test material may be sediments or hydric soil collected from the field or spiked with compounds in the laboratory.

1.2 The test procedure describes a 10-d whole sediment toxicity test with an assessment of mortality and selected sublethal endpoints (that is, body width, body length). The toxicity tests are conducted in 300 to 500-mL chambers containing 100 mL of sediment and 175 mL of overlying water. Overlying water is renewed daily and larval amphibians are fed during the toxicity test once they reach Gosner stage 25 (operculum closure over gills). The test procedure is designed to assess freshwater sediments, however, *R. pipiens* can tolerate mildly saline water (not exceeding about 2500 mg Cl^-/L, equivalent to a salinity of about 4.1 when $Na^=$ is the cation) in 10-d tests, although such tests should always include a concurrent freshwater control. Alternative test durations and sublethal endpoints may be considered based on site-specific needs. Statistical evaluations are conducted to determine whether test materials are significantly more toxic than the laboratory control sediment or a field-collected reference sample(s).

1.3 Where appropriate, this standard has been designed to be consistent with previously developed methods for assessing sediment toxicity to invertebrates (for example, *Hyalella azteca* and *Chironomus dilutus* toxicity tests) described in the United States Environmental Protection Agency (USEPA, (**1**))[2] freshwater sediment testing guidance, Test Methods E 1367 and E 1706, and Guides E 1391, E 1525, E 1611, and E 1688. Tests extending to 10 d or beyond, and including sublethal measurements such as growth, are considered more effective in identifying chronic toxicity and thus delineating areas of moderate contamination (**1-3**).

1.4 Many historical amphibian studies, both water and sediment exposure, have used tests of shorter duration (5 days or less) (for example, **4-7**) and, although both survival and sublethal endpoints were often assessed, there is substantive evidence that tests of longer duration are likely to be more sensitive to some contaminants (**8, 9**). Research performed to develop and validate this test protocol included long-term (through metamorphosis) investigations and other researchers have also conducted long-duration tests with anurans (**7-11**). In the development of these procedures, an attempt was made to balance the needs of a practical assessment with the importance of assessing longer-term effects so that the results will demonstrate the needed accuracy and precision. The most recent sediment toxicity testing protocols for invertebrates have encompassed longer duration studies which allow the measurement of reproductive endpoints (**1, 12**). Such tests, because of increased sensitivity of the sublethal endpoints, may also be helpful in evaluating toxicity. Full life-cycle studies with anurans (including reproduction) are usually not feasible from either a technical or monetary standpoint. However, if site-specific information indicates that the contaminants present are likely to affect other endpoints (including teratogenicity), then the duration of the toxicity test may be increased through metamorphosis or additional sublethal endpoints may be measured (for example, impaired behavior, deformities, time-to-metamorphosis). The possible inclusion of these endpoints and extension of test length should be considered during development of the project or study plan (see 8.1.1).

1.5 The methodology presented in this standard was developed under Department of Defense (DoD) a research program and presented in a guidance manual for risk assessment staff

[1] This guide is under the jurisdiction of ASTM Committee E47 on Biological Effects and Environmental Fate and is the direct responsibility of Subcommittee E47.03 on Sediment Assessment and Toxicology.

Current edition approved Nov. 15, 2007. Published December 2007.

[2] The boldface numbers in parentheses refer to the list of references at the end of this standard.

and state/federal regulators involved in the review and approval of risk assessment work plans and reports (**13**). To develop this method, a number of tests with spiked sediment tests were conducted (**13, 14**). Since development of the methodology it has been used operationally to evaluate field-collected sediments from several state and federal environmental sites (**15, 16**). For most of these studies the preferred test organisms, *Rana pipiens*, was used. At a lead-contaminated state-led site, operated by the Massachusetts Highway Department, *Xenopus laevis* was used in the sediment test system because of availability problems with *Rana pipiens* (**17**), The test method was also used to evaluate sediment toxicity at a cadmium-contaminated USEPA Region 4-led site in Tennessee (**18**). The methodology was used to help characterize potential effects of contaminants on amphibians and to help develop preliminary remedial goals, if warranted. All tests evaluated survival and growth effects after 10 d of exposure in accordance with the methods presented in this standard.

1.6 The use of larval amphibians to assess environmental toxicity is not novel. Researchers have used tadpoles to examine toxicity of metals and organic compounds. Most of these studies have been through water exposure, usually in a manner similar to fish or invertebrate exposure as described in Guide E 729 (**19-29**). Fewer studies have focused on exposure of anuran larvae to sediments, and the methods employed vary widely, from *in situ* enclosures (**30**) to laboratory tests using variable exposure conditions and organism ages (**4, 8, 31-33**). No studies were identified that used the same test conditions as described in this standard. However, several laboratory-based evaluations of sediment effects on amphibians are described in the following subsections.

1.6.1 Sediment toxicity tests conducted in the laboratory with amphibians were performed over a range of test durations from 4 d (**4**, **31**, Guide E 1439-98 Appendix X2) to 12 d (**33**) and through metamorphosis (**8, 32**). Sediment toxicity tests with anurans native to North America were started with larval tadpoles between Gosner stages 23 and 25 (**8, 32, 33**). Test temperatures were between 21 and 23°C and feeding began after tadpoles reached Gosner stage 25. Food sources were Tetramin™ (**8**), boiled romaine lettuce (**32**), or boiled romaine lettuce and dissipated rabbit food pellets (**33**). Tests were conducted in static renewal mode with water replacements conducted at varying rates (daily (**31, 33**), weekly (**8**), every 3 to 5 d (**32**)). Test design (number of replicates, test vessel size, number of organisms per replicate) varied depending on the objective of the study with several tests conducted in aquaria (**32**), large bins (**8**), or swimming pools (**33**). Endpoints evaluated at test termination included survival (**4, 8, 31-33**), growth (**8, 31-33**), bioaccumulation of metals (**8**), developmental rates (**8, 32**), deformities (**31, 32**), swimming speed (**33**) and foraging activity levels (**32**).

1.6.2 To assess the effect of direct contact with the sediments containing PCBs, Savage et al. (**32**) exposed larval tadpoles (Gosner stage 23 to 25; wood frogs (*R. sylvatica*)) to field-collected sediments under conditions that allowed both direct contact with the sediment and separation from the sediment with a 500 µm mesh barrier. The study found that lethal and sublethal effects on tadpoles observed through metamorphosis were more pronounced when direct contact with the sediment was allowed. The test conditions described in this standard allow tadpoles to maintain direct contact with the sediment.

1.6.3 Sediment toxicity testing with the African clawed frog (*Xenopus laevis*) has focused on evaluating the developmental effects of sediment extracts, as opposed to whole sediments, on frog embryos. Methods have been developed which expose blastula stage embryos to sediment by enclosing the embryos in a Teflon mesh insert that rests over the top of the sediment in the sediment–water interface region (**31**, Guide E 1439-98 Appendix X2). These studies are conducted evaluate survival, growth, and physical malformations of the embryos after a 4-d exposure period. The test conditions described in this standard allow more direct contact with the sediment, using older test organisms, and a longer exposure duration.

1.7 Sediment toxicity tests are an effective means for evaluating the impact of sediment contamination on amphibians in a multiple lines of evidence paradigm. The evaluation is most powerful when toxicity testing sampling stations are co-located with sediment analytical chemistry samples and ecological surveys, allowing for a detailed evaluation of the co-occurring data in the ecological risk assessment. The spatial and temporal co-location of toxicity testing and analytical samples is particularly important for establishing contaminant-specific effects and assessing contaminant bioavailability.

1.8 In order for a sediment toxicity test to be sensitive it must be of sufficient duration to measure potential toxicity and it must be conducted during the appropriate developmental stage of the test organism's life cycle. Using recently hatched tadpoles and conducting the sediment exposure test for 10 d to allow the evaluation of growth endpoints meets both of these sensitivity requirements.

1.9 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

2. Referenced Documents

2.1 *ASTM Standards:* [3]

D 4447 Guide for Disposal of Laboratory Chemicals and Samples

E 177 Practice for Use of the Terms Precision and Bias in ASTM Test Methods

E 691 Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method

E 729 Guide for Conducting Acute Toxicity Tests on Test Materials with Fishes, Macroinvertebrates, and Amphibians

E 943 Terminology Relating to Biological Effects and Environmental Fate

E 1367 Test Method for Measuring the Toxicity of

[3] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

Sediment-Associated Contaminants with Estuarine and Marine Invertebrates
E 1391 Guide for Collection, Storage, Characterization, and Manipulation of Sediments for Toxicological Testing and for Selection of Samplers Used to Collect Benthic Invertebrates
E 1439 Guide for Conducting the Frog Embryo Teratogenesis Assay-Xenopus (FETAX)
E 1525 Guide for Designing Biological Tests with Sediments
E 1611 Guide for Conducting Sediment Toxicity Tests with Polychaetous Annelids
E 1688 Guide for Determination of the Bioaccumulation of Sediment-Associated Contaminants by Benthic Invertebrates
E 1706 Test Method for Measuring the Toxicity of Sediment-Associated Contaminants with Freshwater Invertebrates

3. Terminology

3.1 The words "must", "should", "may", "can" and "might" have very specific meanings in this standard. "Must" is used to express an absolute requirement, that is, to state that the design of a test ought to be in a manner that satisfies the specified conditions, unless project goals dictate needed alterations in order to address the study hypotheses. "Should" is used to state that the specified condition is recommended and ought to be met if possible. Although the violation of one "should" is rarely a serious matter, violation of several could render the results questionable. Terms such as "is desirable", "is often desirable" and "might be desirable" are used in association with less important factors, the alteration of which will probably not have substantive effects on test outcome. "May" means "is (are) allowed to," "can" means "is (are) able to" and "might" means "could possibly." In this manner, the classic distinction between "may" and "can" is preserved and "might" is never used as a synonym for either "may" or "can."

3.2 *Definitions*—For definitions of general terms related to toxicity testing and used in this guide, refer to Guide E 943.

3.3 *Definitions of Terms Specific to This Standard:*

3.3.1 *IC25 (25 % inhibition concentration), n*—concentration at which there is a 25 % reduction in organism performance, relative to the control. Performance may be survival or a sublethal measurement such as growth.

3.3.2 *overlying water, n*—water that is placed over the sediment for the duration of the study. Overlying water may be surface water collected from the project site or from a clean lake or reservoir, or may be reconstituted water prepared in the laboratory (for example, moderately hard water; **(34)**).

3.3.3 *reference-toxicant test, n*—a test conducted with a reagent-grade reference chemical to assess the sensitivity of the test organisms. Deviations outside an established normal range may indicate a change in the sensitivity of the test organism population. Reference-toxicity tests are most often performed in the absence of sediment.

3.3.4 *test sediment or test material, n*—sediment that may contain contaminants, which is being evaluated using this test procedure.

4. Summary of Guide

4.1 Each test consists of eight replicates of the test material (for example, field-collected sediment or spiked sediment) and overlying water with five test organisms (recently-hatched tadpoles) per replicate. A laboratory control sediment (sometimes called a negative control) is used to provide *(1)* a measure of the acceptability of the test by indicating the quality of tadpoles, test conditions and handling procedures, and *(2)* a basis for interpreting data from other treatments. The test duration is ten days with an assessment of mortality and selected sublethal endpoints (that is, body width, body length) at the end of the test. Assessments of mortality can be made daily during the test and dead organisms removed. However, similar coloration of the tadpoles and sediment may make it difficult to see the organisms and sediment disturbance should be kept to a minimum. Alternative test durations and sublethal endpoints may be considered based on site-specific needs. The objective of the test is to evaluate whether test materials (spiked or field-collected sediments) are significantly more toxic than the laboratory control or reference sediment(s). Additional evaluations may be performed if an exposure gradient is tested. Statistical evaluations may be conducted to determine whether test materials are significantly more toxic than the laboratory control sediment or field-collected reference sample(s). If the test material is sediment spiked with a known concentration of a chemical stressor or if field-collected sediment contains a measured gradient of a particular chemical of concern, then point estimates (for example, median lethal concentrations (LC50s), 25 % inhibition concentrations (IC25s), or 50 % inhibition concentrations (IC50s)) may be calculated. Field-collected sediments often contain more than one potential chemical stressor and therefore calculating chemical-specific point estimates should only be done with caution. A reference-toxicant test should be run concurrently with a sediment test whenever a new batch or lot of organisms is used.

5. Significance and Use

5.1 While federal criteria and state standards exist that define acute and chronic "safe" levels in the water column, effects levels in the sediment are poorly defined and may be dependent upon numerous modifying factors. Even where USEPA recommended Water Quality Criteria (WQC, **(35)**) are not exceeded by water-borne concentrations, organisms that live in or near the sediment may still be adversely affected **(36)**. Therefore, simply measuring the concentration of a chemical in the sediment or in the water is often insufficient to evaluate its actual environmental toxicity. Concentrations of contaminants in sediment may be much higher than concentrations in overlying water; this is especially true of hydrophobic organic compounds as well as inorganic ions that have a strong affinity for organic ligands and negatively-charged surfaces. Higher chemical concentrations in sediment do not, however, always translate to greater toxicity or bioaccumulation **(37)**, although research also suggests that amending sediment with organic matter actually increases the bioaccumulation of contaminant particles **(38, 39)**. Other factors that can potentially influence

TABLE 1 Advantages and Disadvantages for Use of Sediment Tests (Modified from Test Method E 1706)

Advantages
Measure bioavailable fraction of contaminant(s).
Provide a direct measure of effects on sediment-associated receptors (benthos, larval amphibians), assuming no field adaptation or amelioration of effects.
Limited special equipment is required.
Methods are rapid and inexpensive.
Legal and scientific precedence exist for use; USEPA and ASTM standard methods and guides are available.
Measure unique information relative to chemical analyses or community analyses.
Tests with spiked chemicals provide data on cause-effect relationships.
Sediment-toxicity tests can be applied to all chemicals of concern.
Tests applied to field samples reflect cumulative effects of contaminants and contaminant interactions.
Toxicity tests are amenable to confirmation with natural populations (invertebrate or amphibian surveys).
Disadvantages
Sediment collection, handling, and storage may alter bioavailability.
Spiked sediment may not be representative of field contaminated sediment.
Natural geochemical characteristics of sediment may affect the response of test organisms.
Indigenous animals may be present in field-collected sediments.
Route of exposure may be uncertain and data generated in sediment toxicity tests may be difficult to interpret if factors controlling the bioavailability of contaminants in sediment are unknown.
Tests applied to field samples may not discriminate effects of individual chemicals.
Few comparisons have been made of methods or species.
Only a few chronic methods for measuring sublethal effects have been developed or extensively evaluated.
Laboratory tests have inherent limitations in predicting ecological effects.
Tests do not directly address human health effects.
Motile organisms may be able to avoid prolonged exposure to contaminated media so tests may overestimate actual exposure.
Species used in toxicity testing programs are typically chosen to be representative and protective of the organisms found on-site, but the use of surrogate species cannot precisely predict the health of ecological communities on-site.
Toxicity to organisms in situ may be dependent upon physical characteristics and equilibrium partitioning that are not readily replicated under laboratory conditions.

sediment bioaccumulation and toxicity include pH mineralogical composition, acid-volatile sulfide (AVS) and grain size (**40, 41**). Laboratory toxicity tests provide a direct and effective way to evaluate the effects of sediment contamination on environmental receptors while providing empirical consideration of all of the physical, chemical and biological parameters that may influence toxicity.

5.2 Amphibians are often a major ecosystem component of wetlands around the world, however limited data are available regarding the effects of sediment-bound contaminants to amphibians (**30-32, 41-43**). Laboratory studies such as the procedure described in this standard are one means of directly assessing sediment toxicity to amphibians in order to evaluate potential ecological risks in wetlands.

5.3 Results from sediment testing with this procedure may be useful in developing sediment screening values for amphibians.

5.4 Sediment toxicity test can be used to demonstrate the reaction of test organisms to the specific combination of physical and chemical characteristics in an environmental medium. The bioavailability of chemicals is dependent on a number of factors, which are both site-specific and medium-specific. Although many of these factors can be estimated using equilibrium partitioning techniques, it is difficult to account for all the physical and chemical properties which could potentially affect bioavailability. Sediment toxicity tests may be particularly applicable to evaluating hydrophobic compounds which may not readily partition into the water column. See Table 1 for a summary of advantages and disadvantages associated with sediment toxicity tests.

6. Interferences

6.1 *General Interferences*:

6.1.1 An interference is a characteristic of a sediment or a test system that can potentially affect test organism response aside from those related to sediment-associated contaminants. These interferences can potentially confound interpretation of test results in two ways: *(1)* toxicity is observed in the test sediment when contamination is low or there is more toxicity than expected, and *(2)* no toxicity is observed when contaminants are present at elevated concentrations or there is less toxicity than expected.

6.1.2 These general interferences may include: potential changes in contaminant bioavailability due to manipulation of field-collected sediments during collection, shipping, and storage; the influence of natural physico-chemical characteristics such as sediment texture, grain size, and organic carbon on the response of test organisms; tests conducted with field-collected samples usually cannot discriminate between effects of multiple contaminants. See Guide E 1706 Section 6 for a detailed discussion of several general interferences that pertain to sediment toxicity testing.

6.1.3 Some interferences, such as the presence of indigenous organisms in field-collected sediments, may have less of an impact on toxicity tests conducted with larval amphibians than on tests conducted with sediment invertebrates.

6.2 *Species-Specific Interferences*:

6.2.1 Particular characteristics of individual species that were tested during the development of this method will probably not act as substantial interferences to completion of successful tests. Those species include *Rana pipiens*, *Bufo americanus*, *Rana clamitans*, *Rana palustris*, *Rana sylvatica*, *Hyla chrysoscelis* and *Xenopus laevis*. However, because the sensitivity of these species to all potential sediment-associated contaminants is unknown, use of test organisms for which more toxicity data are available is recommended.

7. Facilities, Equipment, and Supplies

7.1 *Facilities*—While larval amphibians can be acclimated and held for short periods of time in static or static-renewal systems, continuous-renewal/flow-through conditions are preferable shortly after hatching. Tadpoles grow rapidly and, once feeding begins at about Gosner Stage 25 (**44**), ammonia concentrations are likely to increase and oxygen levels may be depressed, making flow-through conditions desirable. Culture/holding tanks and test chambers should be held at a constant temperature, either in an environmental chamber or temperature-controlled water bath. Addition of overlying water in a flow-through system should be gravity-fed from a water source that may be replaced via pumps. Overlying water should be near culture/test temperature although small temperature deviations should have little impact upon test water temperature at the slow rate of water replacement. Low dissolved oxygen concentrations may be remedied by increasing water replacement rates in small increments. If aeration is necessary, air should be free of contaminants including oil, dust and water; a filtration system may be desirable to remove bacterial contaminants. Lighting should be maintained at a 16-h light and 8-h dark cycle unless the test-specific protocol calls for an alternative photoperiod.

7.2 *Special Requirements*—Amphibian eggs and tadpoles can be highly sensitive to alterations in temperature, oxygen deprivation and handling. If eggs are received from an out-of-laboratory source, attention should be paid to how embryos are packed for shipment, shipment time and handling at the laboratory. Shipping containers should be durable, insulated and water tight. Embryos may be contained in large plastic bags sealed with rubber bands. Double bagging is recommended for added security. Oxygenation of the water containing the embryos is recommended before sealing the bags for shipment. Coolers containing embryos should be firmly taped shut before shipment. The use of ice packs or additional insulation in the shipping containers may be needed when outdoor temperatures are elevated or reduced. It is recommended that temperatures be monitored during shipment, if possible, or upon receipt at the laboratory. Upon receipt at the laboratory, eggs should be allowed to hatch with minimal disturbance.

7.3 *Equipment and Supplies*—All equipment used to prepare test sediments or reagents, transfer sediments or organisms and conduct tests, should be decontaminated as outlined below. Table 2 provides a list of the general equipment needed to conduct testing. Glass is the preferable material in which to conduct tests, however, alternative materials such as stainless steel, high-density polyethylene (HDPE), polycarbonate and fluorocarbon plastics may be appropriate, depending upon the contaminants of concern that might be present in the sediment. Used equipment should not be used if there is a possibility of residual contamination that cannot be removed via the washing process. In some cases, test substances present in field-collected sediments or introduced into spiked sediments may not be thoroughly washed from the test vessels. In these cases the test vessels should not be re-used. All new and used equipment needs to be washed in detergent and should be rinsed with dilute acid and deionized water. Rinsing with an organic solvent (for example, acetone) should also be considered for those materials that will not be damaged by the solvent (for example, some plastics) (see Test Method E 1706 section 9.3.6 for a step-by-step cleaning procedure). Materials that should not contact overlying water include copper, cast iron, brass, lead, galvanized metal (that may contain zinc) and natural rubber.

TABLE 2 General Equipment Required for Conducting a 10-d Sediment Toxicity Test with *Rana pipiens*

Stainless steel bowls and spoons or auger to homogenize sediment
Testing chambers (usually 300 to 500 mL beaker with a small-mesh (300 µm) screen covering a hole drilled in the side of the beaker (secured with nontoxic silicone adhesive))
Transfer pipettes
Small nets
Dissecting microscopes
Dissolved oxygen meter and probe
Conductivity meter and probe
pH meter/selection ion meter and probe
Ammonia meter and probe
Reagents and equipment for hardness and alkalinity determinations
Temperature-controlled water bath or environmental chamber capable of controlling to 23 ± 1ºC
Flow-through water delivery system
Buffered 3-aminobenzoic acid ethyl ester, methanesulfonate salt (MS-222 anesthetic) solution.
Food source (TetraMin™)
Appropriate data forms
Metric ruler
Forceps
Statistical software

8. Test Material Collection and Processing

8.1 *Collection*:

8.1.1 Before field collection and preparation of sediments, a sampling/processing procedure should be established that outlines the site- or project-specific steps to be followed. The statistical analyses that will be applied to the data should be considered during the development of the sampling/processing procedure. See Guide E 1391 for additional detail regarding methods for collecting, storing, and characterizing sediment samples.

8.1.2 Sediment should be collected with as little disturbance as possible. It may be desirable to collect sediments from a boat (even if wading is possible) to minimize sediment disruption.

8.1.3 Since the distribution of contaminants in sediment matrices can demonstrate a great deal of spatial variability (**45**), it is desirable to collect multiple replicates from within the delineated study area. At a minimum, multiple samples should be collected and thoroughly composited in the field so the sample better represents environmental conditions.

8.1.4 Large pieces of plant material and other debris, such as large rocks and glass, should be removed and discarded in the field. Alternatively, these materials can be removed in the laboratory prior to test setup.

8.1.5 In general, unless project specific conditions dictate otherwise, sediment should be collected from the top 15 cm of the native horizon, which generally represents the maximum bioactive zone and area of most probable exposure.

8.1.6 The exact collection procedures will depend upon study design. In deeper water where a boat is used, a benthic

grab, dredge or corer should be used (Guide E 1391). At locations where the water is very shallow, including saturated hydric soils, these devices can also be used or a clean trowel or shovel can be used. Whatever collection method is selected, all cleaning and decontamination protocols need to be followed to minimize sample contamination.

8.1.7 The testing procedure described in this standard requires a minimum of about one liter of sediment. Since this amount does not allow for accidental loss, spillage, analytical chemistry, or test reruns, collection of a minimum of two liters is recommended.

8.1.8 The most convenient sample containers are wide-mouth, high-density polyethylene (HDPE) bottles with a screw-on cap. Glass jars may be desirable for some studies where adsorption to plastic surfaces is of concern. However, glass containers require greater care in handling and packing for shipment and are generally more expensive than plastic jars.

8.2 *Storage*:

8.2.1 Light and heat can stimulate and accelerate chemical and biological reactions that may alter chemical composition, promote degradation of potential toxicants, and affect bioavailability. Samples, therefore, should be kept out of sunlight and stored in the dark under refrigeration. Samples should be cooled before shipping, unless the ambient temperature is already <10ºC. Target cooling temperature for sediments is about 4°C (Test Method E 1367). Ice or blue ice should be included with the samples when they are shipped. Samples should not be frozen as freezing can alter sediment characteristics.

8.2.2 For additional information on sediment collection and shipment see Guide E 1391.

8.2.3 It is desirable to initiate tests as soon as possible following field collection of sediments (Test Method E 1706). Several studies have addressed the question of storage time for sediments, and the conclusions reached in these studies vary considerably. Where the potential chemical stressors are known to recalcitrant, storage under the conditions described in 7.9 should allow the sample to remain stable for longer periods. However, some labile chemicals (for example, ammonia and volatile organics) can degrade or volatize during storage. For these labile materials, a maximum holding time of two weeks (from the time of sample collection to test initiation) is recommended **(46)**. However, more stable sediments can be stored for much longer periods of time with little change in toxicity.

8.2.4 During even short periods of storage, density differences will results in settling in samples, resulting in a heterogeneous mixture. Therefore, prior to test initiation, the sediment should be homogenized again, even if it was already mixed in the field. In most situations, overlying water should not be drained off the sample, but should be remixed with solid material. If, after 24 hours of undisturbed settling, >75 % of the sample volume can still be considered standing water, it may be desirable to remove some or all of that water so as to ensure that the test material will be a solid matrix.

8.3 *Manipulation*:

8.3.1 *Homogenization*:

8.3.1.1 Homogenization can be accomplished by using a tumbling or rolling mixer or other suitable apparatus. It can also be done using a stainless steel auger and drill or simply by hand with a stainless steel spoon. A minimum interval (at least three minutes) should be established for mixing each sample. A more heterogeneous sample would indicate the need for a longer mixing time. Additional large debris should be removed at this time. Sieving of samples is not recommended, however, indigenous organisms can be removed by hand during the mixing process. Special attention should be paid to any predaceous organisms that might be present in the collected sample. Augers, spoons, and any other equipment that comes in contact with the sediment during homogenization must be washed and decontaminated between samples.

8.3.2 *Sediment Spiking*:

8.3.2.1 Test sediment can be prepared by manipulating the properties of a control sediment (Test Method E 1706). Mixing time **(45)** and aging **(47)** of spike sediment can affect bioavailability of chemicals. If tests are initiated within only a few days of spiking a sediment, the spiked chemicals may not be at equilibrium with the sediment. There are not, however, specified equilibrium intervals for all chemicals that might be spiked into sediment. Such specifications would not be reasonable since sediment characteristics will play a major role in time to equilibration as well as equilibration concentrations. For a series of spiked sediment studies, where results will be compared, spiking methods should be consistent and the amount of time between spiking and test initiation should also be consistent.

8.3.2.2 The test material(s) should be at least reagent grade, unless a test using a formulated commercial product, technical-grade or use-grade material is specifically needed. Before a test is initiated, the following should be known about the test material (not all of this information may be available): *(1)* the identity and concentration of major ingredients and impurities, *(2)* solubility in test water and water used to prepare any stock solutions, *(3)* log K_{ow}, BCF, persistence, hydrolysis and photolysis rates, *(4)* estimated toxicity to the test organism, *(5)* toxicity to humans and potential handling hazards, *(6)* if and when analytical samples will be collected, how much material will be needed to obtain the needed resolution and preservation methods, and *(7)* recommended handling and disposal methods.

8.3.2.3 Different sediment spiking methods are available. Sediment spiking techniques used during development and validation of the amphibian sediment test method **(13)** were previously employed for incorporation of both inorganic contaminants and organic chemicals into sediment **(42)**. The procedure included: *(1)* place appropriate (considering testing and analytical needs) amount of sediment in a mixing jar, *(2)* if sediment is dry, wet it with deionized water to ensure holes in the sediment will remain open, *(3)* using a 10-mL or 5-mL pipet, punch at least five holes into the sediment to different depths, *(4)* distribute equally to each hole the volume of the stock solution needed to achieve the desired target concentration of test material. The stock solution may be an inorganic salt dissolved in water (for example, copper as $CuCl_2$). If a hydrophobic chemical is to be tested, it may first be dissolved

into a stock solution using a carrier solvent (for example, acetone or methanol). A surfactant should not be used in the preparation of a stock solution because it might affect the bioavailability, form or toxicity of the test material. If a carrier solvent is used, a solvent control must also be prepared which contains the solvent but not the contaminant to be tested. See USEPA (**48**), Guide E 1391, and Test Method E 1706 for additional details regarding sediment spiking techniques.

8.3.2.4 Once spiked, the sediments need to be thoroughly mixed to incorporate the chemical into the sediment and create a homogenized matrix. Homogenization methods include roller mixers, end-over-end mixers stainless steel kitchen mixers, mixing manually with a spoon or a combination of these. Mixing times, speeds and temperatures should be consistent among treatments, replicates and tests.

8.3.3 *Test Concentration(s) for Laboratory-Spiked Sediments*:

8.3.3.1 If a test is intended to generate an LC50, IC50 or IC25 of a test chemical, a concentration series should be created that will bracket that effect concentration. If mortality is one of the desired endpoints, at least one test concentration should produce greater than 50 % mortality and there should be two or more concentrations with partial mortality. Determining the concentration(s) that will result in desired lethal or sublethal effects can be difficult if *(1)* the environmental toxicity of the test material is unknown and/or *(2)* the impact(s) of sediment characteristics is/are unknown. The latter can be particularly important since there are many factors that can significantly affect toxicity (**37-41**). It may be desirable to conduct a range-finding test in which the organisms are exposed to a control and three or more concentrations of the test material that differ by a factor of ten. For example, test concentrations in a range-finding test may include the control, 10, 100 and 1000 mg/kg.

8.4 *Sediment Characterization*:

8.4.1 It is recommended that a subsample of each field-collected or spiked sediment be analyzed for at least the following parameters: pH, total organic carbon (TOC), particle size distribution (percent sand, silt, clay). Similar analyses should also be conducted on laboratory control sediment and reference sediment(s).

8.4.2 Further characterization may be warranted depending on the objectives of the study. This may include chemical analyses of inorganic and organic compounds of interest, ammonia, pore water chemistry, chemical oxygen demand, sediment oxygen demand, oxidation-reduction potential (Eh), acid volatile sulfides (AVS), and simultaneously extracted metals (SEM), or other analyses depending on the program.

8.4.3 Chemical and physical data should be obtained using appropriate standard methods whenever possible. For those measurements for which standard methods do not exist or are not sensitive enough, methods should be obtained from other reliable sources.

8.4.4 Sediment characterization helps to evaluate sediment homogenization and accuracy of sediment-spiking, and identifies potential chemical or physical stressors for test organisms.

9. Test Organisms

9.1 *Species*—Test organisms are recently hatched tadpoles of small North American anurans. The preferred species is the Northern Leopard Frog, *R. pipiens*. Sediment toxicity testing conducted with both *R. pipiens* and the American toad, *B. americanus*, during the development of this standard indicated that *R. pipiens* was generally more sensitive to spiked sediments containing metals (cadmium, copper, lead, or zinc) than was *B. americanus* (**13**). A review of amphibian data presented in U.S. EPA ambient water quality criteria documents for cadmium, copper, and zinc (**13**) and relative sensitivity data evaluating amphibian aquatic LC50s (**49**) indicate that *R. pipiens* is considered to be sensitive to metals, relative to other frog, toad, and salamander species. Other ranid species (*R. catesbeiana*, *R. palustris*) were also sensitive to the metals reviewed (**13, 49**). The potential for field-collection of *R. pipiens* eggs with minimal impact to local communities was also a consideration in the selection of this species as the preferred test species. Other species may be used for testing if handling and holding conditions are known.

9.2 *Sources*—While adults of several species of toads and frogs are available for most of the year from commercial suppliers of living organisms, availability of eggs is more limited. Eggs of *R. pipiens* can be collected in the wild during the spring. Since it may be difficult to distinguish between the eggs of related anuran species, collectors should be well-trained in species' habitats and identification. Collectors should comply with all state and federal regulations and be in possession of current collecting permits, if required. If possible, adult animals should also be collected for identification in the same area that eggs are being collected.

9.2.1 Eggs of *R. pipiens* can be obtained from commercial suppliers or be field collected from about November until April. Eggs that are produced and fertilized in the laboratory are preferable since the taxonomy is known. Researchers are encouraged to use available resources to find suppliers.

9.3 *Care and Handling*—Eggs received from commercial suppliers or collected in the wild should be subjected to a minimum of handling. Suppliers generally package and ship eggs in sealed bags or other containers that have been injected with oxygen (dissolved oxygen levels should be maintained above 4 mg/L to avoid stressing the test organisms). Hatching success is higher if handling of eggs is minimized; if possible eggs should left in the original shipping package until development is verified and organisms are near hatching stage. Upon receipt, bags containing eggs should be allowed to slowly rise (no more than 3°C per hour) to test temperature (avoid rapid temperature changes). If eggs arrive in containers that have not been injected with oxygen or otherwise cannot be left intact, organisms should be transferred to an aquarium or other holding container and slowly brought to test temperature.

9.3.1 Time to hatch will depend upon age at the time of shipping. Once the young embryos have developed into a recognizable tadpole and are actively moving, the bag can be opened and the eggs/early stage tadpoles placed in an aquarium or other large chamber.

9.3.2 Once the eggs/tadpoles are released for the shipping container to an aquarium or other chamber, shipping water

TABLE 3 Developmental Stages of Anuran Embryos (from Gosner (44) and Shumway (51))

Stage	Approximate Age at 18ºC (h) for Stages 1 through 25	Major Characteristics/Formations of the Stage
1	0	Prior to fertilization
2	1	Appearance of post-fertilization gray crescent
3	3.5	Two blastomeres
4	4.5	Four blastomeres
5	5.7	Eight blastomeres
6	6.5	Sixteen blastomeres
7	7.5	Thirty-two blastomeres
8	16	Mid-cleavage
9	21	Late cleavage
10	26	Appearance of dorsal lip of blastopore
11	34	Mid-gastrula, blastoporal lip invaginating along semicircle
12	42	Late gastrula, blastoporal lip invaginating around the circular yolk plug. Yolk plug diameter ~ ⅕ diameter of gastrula
13	50	Neural plate, blastopore forming slit
14	62	Neural folds
15	67	Rotation of embryo
16	72	Neural tube
17	84	Tail Bud
18	96	"Tadpole" shape becoming distinct; muscular response to stimulation
19	118	Heart beat; external gill buds; hatching begins
20	140	Complete hatching; swimming upon physical stimulation; capillary circulation in first gill
21	162	Mouth open; transparent cornea; tail length approximately equal to length of head and body
22	192	Transparent epidermis; capillary circulation in tail; asymmetrical appearance from dorsal aspect; left gills filaments more apparent
23	216	Opercular fold apparent; asymmetrical from ventral aspect
24	240	Operculum covering right external gills; external gills on left side still apparent; sucker represented by two small prominences
25	284	Operculum complete; no external gill filaments; Sucker represented by two pigmented patches; begin feeding; gut clearly visible
26–30		Hind limb buds appear and grow progressively larger; spiracle present on left side (most North American tadpoles)
31		Toes begin to develop on hind limbs
32–37		Toes on hind limbs grow progressively distinct; all five toes apparent at stage 37
38–40		Toes continue to lengthen; metatarsal and subarticular tubercles develop
41		Tail begins to shorten; cloacal tail piece disappears; skin over forelimbs becomes transparent; lateral forelimb "bulges" appear
42–45		Forelimbs break through membrane; Face shortens; mouth lengthens; posterior edge of mouth extends beyond posterior edge of eye; tail absorption continues
46		Metamorphosis complete; tail stub usually present; froglets must have physical platform to leave the water

should be slowly replaced with culture/overlying water. This should be done by initially adding culture/overlying water at a proportion of no more than 10 % for one hour. If organisms do not appear to be adversely affected, increase the amount of culture/overlying water by about 15 to 25 %/ hour for 4 to 5 hours.

9.3.3 Additional acclimation of test organisms should not be needed under most circumstances.

9.3.4 Low dissolved oxygen will increase organism stress and may cause mortality in the holding chamber or result in increased mortality during a test. Dissolved oxygen should not be allowed to fall below 3.0 mg/L. If needed, gentle aeration should be initiated using a small pipette and low bubble rate.

9.3.5 Always wear laboratory gloves (for example, latex; talc-free) when handling eggs. Direct contact with eggs or tadpoles should be avoided to minimize stress on the organisms. Transfer eggs and tadpoles gently and with minimal handling time.

9.4 Once embryos have reached a distinctive tadpole shape (about Gosner stage 19-20) they are far less prone to mortality from handling.

9.5 A sub-sample of specimens should be collected and preserved for species verification.

10. Hazards

10.1 Some test materials, as well as some materials used to preserve test organisms, may be inherently hazardous. Caution needs to be used when handling these materials. Guidelines for the handling and disposal of hazardous materials should be strictly followed (Guide D 4447). When working with any potentially hazardous materials, including those used for analytical measurements (for example, acid used during alkalinity titrations), users need to wear appropriate protective equipment (for example, safety glasses and gloves). Common laboratory protective wear should also be used to reduce exposure to potential biological hazards (for example *Salmonella*, *Vibrio* ssp.). All laboratory-specific health and safety considerations should be followed. (see Test Method E 1706 for additional detail).

11. Procedure

11.1 *Experimental Design*—Each test consists of eight replicates of the test material (e.g., field-collected sediment or spiked sediment) and overlying water with five test organisms (recently-hatched tadpoles) per replicate. It may be necessary to make modifications of the basic experimental design to

accommodate project-specific circumstances, including shortage of available test sediment (for example,, scarce depositional areas in riverine systems), bioaccumulation (need for extra tissue) or additional analytical measurements. A laboratory control sediment (negative control) must be included with all tests and reference sediment(s) may be included when field-collected sediments are tested.

11.1.1 A laboratory control sediment is a sediment that is essentially free of contaminants and is used to ensure that contamination is not introduced during the experimental set up and that test organisms are healthy. This sediment is not necessarily collected near the site of concern. A reference sediment is collected near an area of concern and is used to assess sediment conditions exclusive of material(s) of interest. Testing a reference sediment provides a site-specific basis for evaluating toxicity.

11.2 *Initiating a Test*:

11.2.1 *Adding Sediment to Test Chambers*—The day before the test is to start (Day -1) sediment should be thoroughly homogenized and 100 mL of sediment is added to each test chamber. Overlying water (175 mL) is added to each test chamber in a manner that minimizes disturbance of the sediment. This can most easily be accomplished by pouring against the inside of the chamber. The sediment should be left undisturbed overnight.

11.2.1.1 On the day of test setup (Day 0), withdraw an adequate amount of overlying water from each treatment to conduct all necessary chemical characterizations and analyses. Removal of water should be done with as little sediment disturbance as possible. At a minimum, dissolved oxygen, temperature, pH, conductivity, hardness, alkalinity and ammonia should be measured in each treatment. If samples are collected for other parameters, such as metals, then proper handling and preservatives should be used (see Guide E 1391 for additional detail).

11.2.1.2 Overlying water should be renewed during a test, unless nonrenewal is a fundamental part of the test design. Renewal may be done continuously through a water-delivery system, including diluters or drip-manifolds, or by static replacement. In either case, the volume of water addition in a 24-hour period should not exceed 2 to 3 volumes of overlying water (about 350 to 525 mL). A water-delivery system should be calibrated at test initiation and examined on a daily basis so all test chambers receive about the same amount of water. If manual water addition is conducted, no more than 80 % of the overlying water should be removed at any one time and sediment disturbance should be minimized. The toxicity test is designed to include both sediment and water column exposure to contaminants so it is important to maintain the indicated renewal rates in order to avoid excessive dilution of water column constituents that could lead to an underestimation of sediment toxicity.

11.2.2 *Addition of Test Organisms*—Test organisms should be handled as little as possible. Organisms should be added to the overlying water using a pipette with a large enough bore to prevent constriction and damage to the animals. The animals should be gently released just below the water's surface. The developmental stage (Gosner stage) of the tadpoles should be documented by examining a subset of at least 10 organisms.

11.2.2.1 Development stage should be determined in accordance with descriptions provided by Gosner (**44**). Table 3 provides a summary of the major characteristics of each stage between fertilization and metamorphosis.

11.3 *Monitoring a Test*—All chambers should be checked daily for dead organisms and behavior. Tadpole coloration often makes it difficult to see them against sediment, however, if dead organisms are found, they should be removed with a pipette. Animals that die during a test need only be kept if sublethal observations are to be made or tissue will be analyzed for chemicals of concern. Organisms need to be preserved appropriately for the analyses (see Guide E 1688 for additional detail). The overlying water renewal system should be checked daily to ensure adequate flow and an acceptable addition rate. Screens on the outside of test chambers should be checked

TABLE 4 Test Conditions for Conducting a 10-d Sediment Toxicity Test with *Rana pipiens*

Parameter	Conditions
1. Test type:	Whole-sediment toxicity test with renewal of overlying water
2. Temperature:	23 ± 1°C
3. Light quality:	Wide-spectrum fluorescent lights
4. Illuminance:	About 100 to 1000 lux
5. Photoperiod:	16L:8D
6. Test chamber:	400 to 500-mL glass or plastic beaker or chamber with drainage system
7. Sediment volume:	100 mL
8. Overlying water volume:	175 mL
9. Renewal of overlying water:	Continuous flow-through of overlying water or daily static water addition (not to exceed 2 to 3 volume additions/day)
10. Age of organisms:	≤72 hours, 24 hours or less preferred at the start of the test
11. Number of organisms/chamber:	5
12. Number of replicate chambers/treatment:	Depends on the objective of the test. Eight replicates are recommended for routine testing (see 11.1)
13. Feeding:	4 mg of ground TetraMin™ per vessel daily after tadpoles reach stage 25; reduced proportionally with mortality
14. Aeration:	None, unless dissolved oxygen in overlying water drops below 3.0 mg/L.
15. Overlying water:	Site water, site water match (hardness and alkalinity), natural lake or groundwater, or reconstituted laboratory water (for example, U.S. EPA moderately hard **(5)**)
16. Test chamber cleaning:	If screens become clogged during a test, gently brush the *outside* of the screen
17. Overlying water quality:	Hardness, alkalinity, conductivity, pH, dissolved oxygen, and ammonia at the beginning and end of a test. Temperature and dissolved oxygen daily. Ammonia may also be measured periodically (Days 1, 3, and 7).
18. Test duration:	10 d
19. Endpoints:	Survival and growth
20. Test acceptability:	Minimum mean control survival of 80 %; mean body width of at least 4 mm and body length of at least 7 mm for test organisms in the control sediment. See Table 6 for additional performance-based criteria.

daily to ensure that water is adequately draining. Clogged screens can be brushed to remove impinged debris; cleaning and brushing should only be done with a small, clean brush, cleaning tool or gloved finger. Test conditions are summarized in Table 4 and a list of daily activities is presented in Table 5.

11.3.1 *Monitoring of Overlying Water Characteristics*—Conductivity, hardness, alkalinity, pH and dissolved oxygen must be measured in all treatments at the beginning and end of the test. Dissolved oxygen should also be measured daily. Temperature should be measured continuously in the environmental chamber or water bath and periodically in each treatment (for example, days 3, 6 and 9). If continuous temperature monitoring is not available then instantaneous temperature in each treatment should be measured daily. In any test chamber where mortality has occurred, dissolved oxygen and pH should be measured on the day when mortality was observed.

11.3.1.1 If dissolved oxygen in any one chamber of a treatment is less than 3.0 mg/L, then dissolved oxygen in other chambers within that treatment should be checked. The flow rate (drip rate if a continuous drip manifold is used) in any one chamber can be increased slightly to increase dissolved oxygen. All test chambers should be treated the same relative to test condition modifications (for example, increase in water delivery rate). If after one hour, dissolved oxygen is still <3.0 mg/L, then all of the test chambers within that treatment should be aerated. Set aeration tubes or pipettes so that the narrow tip is submerged not more than 0.5 cm. Bubble rate should be slow and should not disturb the sediment or overly agitate the water's surface to avoid the release of volatile substances. Occasional dissolved oxygen measurements of <3.0 mg/L during a test is not sufficient reason to discard test data, although evidence of extended oxygen depression should be considered with regard to possible adverse affects.

11.3.1.2 Ammonia should be measured in the overlying water on Day 0, at test termination and periodically during the test, for example, days 1, 3 and 7. If ammonia concentrations are >5.0 mg/L (NH_3-N) in any treatment, than a second sample should be collected and measured from another replicate. Tadpoles are sensitive to elevated ammonia, although *R. pipiens* has been found to be less sensitive than some other anurans (**7, 48**). Elevated ammonia concentrations may be a reflection of sediment characteristics and should be taken into consideration when interpreting test results. Test specifications are listed in Table 4.

11.3.1.3 *Temperature*—Target test temperature is 23 ± 1ºC. Daily mean temperature (directly in the water bath or a surrogate test chamber in the water bath or environmental chamber) should be within 1ºC of 23ºC; instantaneous temperature should be 23 ± 3ºC. Continuous monitoring of bath or environmental chamber temperature is preferred.

11.3.2 *Feeding*—Feeding should begin when tadpoles reach Gosner stage 25 (**44**), that is, when an operculum develops and external gills disappear. About 3 to 4 mg of ground, dry TetraMin™ is added daily to each test chamber. Adding excess food should be avoided since it can cause dissolved oxygen depression and may also affect the toxicity of certain chemicals (**39**). Tadpoles in at least three chambers should be examined daily to determine if stage 25 has been reached (see Table 3 or (**44**)). Some toxicants may delay development; feeding of organisms may start on different days for different treatments. It takes about 3 to 5 days for newly-hatched tadpoles to reach stage 25. If older organisms were used, feeding will begin sooner. The amount of food added to each chamber should be decreased if some animals have died. In general, follow the USEPA (**34**) procedures for conducting short-term chronic tests with fathead minnows, *Pimephales promelas*. That is, if 50 % or more of the test organisms have died in a test chamber, reduce the amount of food by 50 %.

11.4 *Ending a Test*—Final water characterization measurements should be made and live organisms should be removed from each chamber with a pipette. All live organisms from a replicate chamber should be placed into a separate, small glass or plastic beaker or cup containing 10 to 20 mL of clean (unchlorinated) water (for example, USEPA Moderately Hard Water (see (**5**) or Guide E 729)). All chambers should be carefully examined for any missing organisms. Dead tadpoles will decompose rapidly and may easily blend into sediment. Unaccounted-for organisms should be considered mortalities.

11.4.1 *Sublethal Measurements*—Live tadpoles should be anesthetized or euthanized before sublethal measurements are

TABLE 5 General Activity Schedule for Conducting a 10-d Sediment Toxicity Test with *Rana pipiens*

Day	Activity
-1	Add homogenized sediment into each test chamber, place chambers into exposure system, and add overlying water.
0	Begin flow through system or conduct first water replacement if using static renewal. After at least one hour collect overlying water for initial water characterization (hardness, alkalinity, conductivity, pH, dissolved oxygen, and ammonia, and total residual chlorine). Add 5 tadpoles to each test chamber. Release organisms under the surface of the water. Archive and preserve 5 to 10 organisms for possible examination of metamorphic stage.
1 to 9	Measure temperature, dissolved oxygen. Measure ammonia periodically in each treatment during the toxicity test (for example, Days 1, 3, and 7). Observe behavior and metamorphic stage of test organisms. Remove dead organisms. Feed 4 mg of ground, dry TetraMin™ per chamber daily after tadpoles reach Gosner stage 25.
10	Measure temperature, dissolved oxygen, pH, conductivity. Collect samples for final water quality measurements (for example, hardness, alkalinity, ammonia), as indicated in project requirements. Remove and count live organisms from each test chamber and transfer them to small beakers (glass or plastic) containing 10 to 20 mL of clean (unchlorinated) water. Euthanize or anesthetize test organisms prior to making sublethal measurements. Measure the maximum body width and body length (snout-to-vent length).

made. The use of a buffered 3-aminobenzoic acid ethyl ester (MS-222) solution is recommended. To each of the small beakers or cups containing live tadpoles, add about 1 mL of a MS-222 stock solution (2 g/L) buffered to about pH 7 using an appropriate buffer medium (for example, sodium bicarbonate). If organisms continue to move after several minutes, add a few more drops of the MS-222. Tadpoles should not be left for an extended period of time in the MS-222 solution as it may cause disintegration of tissue.

11.4.1.1 Using a metric ruler, measure the maximum body length along the center line of the body, excluding the tail (snout–to-vent length). Also, measure the maximum body width. Do not push down on the tadpole body as that will distort these measurements.

11.4.2 Digital photographs and digitizing software may also be used to quantify sublethal measurements.

11.4.3 Statistical evaluations for lethal and sublethal endpoints may be conducted using comparisons to results from the laboratory control or a field-collected reference sample(s). If the test was one in which sediment was spiked with a hydrophobic test material dissolved in a solvent carrier and a solvent control was included in addition to a laboratory control sediment, then survival and growth should be compared between the two controls. If a statistically significant difference is detected between the controls, then only the solvent control may be used for meeting the acceptability of the test and as the basis for calculation of results. The laboratory control may provide additional information on the general health of the test organisms. If no statistically significant difference is detected between the controls, the data from both controls may be pooled and used as a basis for meeting acceptability criteria and as a basis for calculation of results. If the solvent control is markedly different from the laboratory control, it is possible that the data are compromised by experimental artifacts and may not accurately reflect the toxicity of the test material in natural sediments. In such circumstances, the test may need to be repeated or alternative means of test material introduction explored. A discussion of possible statistical evaluations is presented in Appendix X2 but may be modified based on project-specific requirements.

11.5 *Studies Conducted Beyond Ten Days*—If site-specific information indicates that longer duration toxicity tests should be conducted, the daily activities described previously should be followed until test termination.

11.5.1 Activities conducted at test termination will be similar to those conducted for the 10-d toxicity test but may also include inspection for deformities, observations of impaired behavior (prior to anesthetizing), or developmental stage. Feeding should be increased in proportion to the increase in body size of the test organisms. If growth is not affected, the amount of food can be increased by about 2 mg per chamber every five days; not to exceed 12 mg per chamber. If the growth of organisms is diminished, feeding levels should remain unchanged or be increased at a slower rate. Excess food on the surface, sediment or sides of the test chambers indicates that too much is being added and the amount of food should be reduced. At metamorphosis, most larval anurans stop eating as their internal and external physiology undergoes substantial alterations in the shift from a fully aquatic tadpole to an amphibious adult (**43**). As the organisms within a replicate approach late-stage metamorphosis, the amount of food consumed will drop substantially and feeding amounts should proportionally decrease to initial levels or less. At some point, if no feeding behavior is observed and unconsumed food is present, feeding may be stopped within a particular replicate.

11.5.2 If the toxicity test is to be conducted through metamorphosis, some modifications would need to be made in the test system. At complete metamorphosis (about Gosner stages 45 and 46) froglets crawl out of the water. Failure to provide a means of leaving the water will result in tadpole death. Providing an "emergence platform" may be difficult if the original test chambers were beakers or similar vessels. Sediment, water and organisms can be transferred to a vessel with a larger surface area that provides better access for the researcher (for example, a 12 by 25 cm plastic chamber). The emergence platform can be constructed in several ways, but the froglet will need to be able to crawl from the water to air. Possible emergence structures include inclined glass or plexiglass, bricks or stones, sponges and arched pieces of heavy, nylon netting. Any material used as an emergence structure needs to be decontaminated as outlined in 7.3 and should not block water circulation or prevent tadpoles from moving freely about the test chamber.

11.6 *Reference Toxicant Testing*—Reference toxicant tests involve exposing organisms that are used to start a sediment study to known concentrations of a specific reagent-grade reference chemical in water-only exposures in order to assess their sensitivity to a toxicant challenge. Organisms of a given species should demonstrate a consistent response to a reference toxicant. Deviations outside an established normal range may indicate a change in the sensitivity of the test organism. A reference toxicant test must be conducted with each new lot or batch of test organisms that are used to initiate a test. Test conditions for conducting reference-toxicity tests with *R. pipiens* are outlined in Table 6. The procedures can also be used for conducting reference-toxicity tests with the test organisms outlined in Appendix X1.

11.6.1 There are several chemicals that are used as reference toxicants. Copper chloride ($CuCl_2$) has been found to produce consistent responses from the test organisms when organism age and test water are held constant. Other possible reference toxicants include salts such as NaCl and KCl. A reference-toxicant concentration series should be selected that will provide partial mortalities at two or more concentrations of the test chemical in order to allow calculation of appropriate point estimates (LC50, EC50).

11.6.2 A reference toxicant control chart should be prepared for each toxicant (if difference ones are used) that progressively illustrates reference toxicant test results. Results should be illustrated as the calculated value for a test, bracketed by the upper and lower control limits. The control chart should include the 20 most recent reference toxicant data points (**34**).

11.6.3 If the reference toxicity results from a given study fall outside the "expected" range (more than 2 standard deviations), the sensitivity of the organisms and the acceptability of the study may be in question. However, at a 0.05

TABLE 6 Recommended Test Conditions for Conducting Reference-Toxicity Tests

Parameter	Conditions
1. Test type:	Water-only test
2. Dilution series:	Control and at least 5 test concentrations (0.5 dilution factor)
3. Toxicant:	KCl, NaCl, or $CuCl_2$
4. Temperature:	23 ± 1°C
5. Light quality:	Wide-spectrum fluorescent lights
6. Illuminance:	About 100 to 1000 lux
7. Photoperiod:	16L:8D
8. Renewal of water:	At least every 48 hours
9. Age of organisms:	≤72 hours, ≤24 hours preferred
10. Test chamber:	250-500 mL glass or plastic beaker
11. Volume of water:	100 mL (minimum)
12. Number of organisms/chamber:	5
13. Number of replicate chambers/treatment:	3 minimum
14. Feeding:	4 mg/day to each test chamber when organisms reach Gosner stage 25
15. Substrate:	None
16. Aeration:	None, unless DO ≤ 3 mg/L
17. Dilution water:	Culture water, well water, surface water, site water, or reconstituted laboratory water (for example, USEPA moderately hard **(5)**)
18. Test chamber cleaning:	None
19. Water quality:	Hardness, alkalinity, conductivity, dissolved oxygen, and pH at the beginning and end of a test. Temperature daily.
20. Test duration:	7 d
21. Endpoint:	Survival (LC50) and growth (IC25)
22. Test acceptability:	80 % control survival

TABLE 7 Test Acceptability Requirements for a 10-d Sediment Toxicity Test with *Rana pipiens*

A. It is recommended for conducting a 10-d test with *Rana pipiens* that the following performance criteria be met:
1. Age of *R. pipiens* at the start of the test must be ≤72 hours.
2. Average survival of *R. pipiens* in the control sediment must be greater than or equal to 80 % at the end of the test. Growth of test organisms should be measurable in the control sediment at the end of the 10-d test (mean body width of at least 4 mm and body length of at least 7 mm for test organisms in the control sediment).
3. Hardness, alkalinity, and ammonia of overlying water typically should not vary by more than 50 % during the test, and dissolved oxygen should be maintained above 3.0 mg/L in the overlying water.
B. Performance-based criteria for maintaining *R. pipiens* include the following:
1. It may be desirable for laboratories to periodically perform water-only reference toxicity tests to assess the sensitivity of culture organisms (see 11.6). Data from these reference toxicity tests could be used to assess genetic strain or life-stage sensitivity of test organisms to select chemicals.
2. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.
3. Physiological measurements such as lipid content might provide useful information regarding the health of the cultures.
C. Additional requirements:
1. All organisms in a test must be from the same source.
2. Storage of sediments collected from the field should follow guidance outlined in 8.2.
3. All test chambers should be identical and should contain the same amount of sediment and overlying water.
4. Negative-control sediment and appropriate solvent controls must be included in a test. The concentration of solvent used must not adversely affect test organisms.
5. Test organisms must be cultured and tested at 23°C (±1°C).
6. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.
7. Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

probability level, it is expected that, by chance alone, one in 20 test results would fall outside the control limits. If more than one in 20 reference toxicant tests fall outside of the control limits, the laboratory should investigate possible sources of variability and take corrective action, if appropriate. If serious problems are not found, then associated test results may be considered acceptable.

12. Acceptability of Test

12.1 Acceptable survival in the test control is 80 % or greater. Control organisms (*R. pipiens*) should also have a mean body width of at least 4 mm and a body length (snout-to-vent) of 7 mm. If alternative test species are used , researchers may need to complete appropriate data gathering tests to determine acceptable size criteria prior to conducting the toxicity tests. If control performance does not meet these criteria, then the test data should be examined to determine if it is acceptable. Test acceptability criteria are presented in Table 7. Even if control performance does not meet these criteria, test data may still be valuable and yield important results. The following test data should be examined:

12.1.1 Survival in all test treatments. If survival in all test treatments is greater than in the control, then statistical evaluations of test sediments against the laboratory control do not need to be conducted. Statistical comparisons against the reference sediments may still be conducted.

12.1.1.1 If poor performance is observed in the laboratory control, such studies should be repeated to ensure accurate results. However, the scope or sampling associated with some studies may make it difficult or impossible to repeat a study. There may be cases where performance in the negative control is poor, but performance criteria are met in reference sediment included in the study design. In these cases, it might be reasonable to infer that other samples that demonstrate organism performance equivalent to, or better than, the reference

sediment are probably not toxic; however, any samples showing poor performance should not be judged to be toxic, since it is unknown whether the factors that caused poor control performance might have also caused poor performance in the test treatments.

12.1.2 Variability within a treatment. If mortality is highly variable and scattered throughout the test, then the test might not be acceptable. Highly variable survival may be due to variations in water chemistry (for example, low dissolved oxygen or elevated ammonia due to excess food in some chambers), variability in organism health, or differences in how chambers were treated (for example, different amounts of food or flow rates of overlying water).

12.2 There are no specific acceptability requirements for survival in test treatments collected from reference stations. If reference sediment was collected and if survival in the reference sediments is significantly reduced, then questions are raised as to the appropriateness of the reference site.

12.3 Reference toxicant data for a given batch of organisms should fall within the historical 95 % limits for that species. However, data falling outside the range does not necessarily indicate automatic rejection of the data.

13. Report

13.1 Report the following information:

13.1.1 Identity of the test material (for example, test sediments and reference sediment, if collected), investigator(s) name, location of laboratory, and dates of test initiation and termination.

13.1.2 Source of test material (if a specific chemical or compound), its lot number, composition (identities and purity), known physical and chemical properties and the identity and source of any solvent used.

13.1.3 Source of the laboratory control sediment and overlying water.

13.1.4 Chemical characteristics of test material, laboratory control sediment, and overlying water, if available.

13.1.5 Source of test organisms, scientific name (and subspecies, if appropriate), life stage, treatments, acclimation procedures and food.

13.1.6 Description of the experimental design, test chambers or compartments, amount of sediment and overlying water, replicates, organisms per replicate, lighting, food type and feeding rate.

13.1.7 Range of measured concentrations of dissolved oxygen, temperature, pH and conductivity of overlying water.

13.1.8 Chemical and biological monitoring information recorded on daily data sheets during the toxicity test.

13.1.9 A table that lists the percent mortality and mean sublethal results (that is, body width, body length) for each test material.

13.1.10 The names of the statistical tests employed, the alpha-levels of the tests, and some measure of the variability of the hypothesis tested.

13.1.11 Anything unusual about the test and any deviations from the test-specific protocol or procedures followed.

14. Precision and Bias

14.1 *Determining Precision and Bias*—Precision is a term that describes the degree to which data generated from replicate measurements differ and reflects the closeness of agreement between randomly selected test results. Bias is the difference between the value of the measured data and the true value and is the closeness of agreement between an observed value and an accepted reference value (Practices E 177 and E 691). Quantitative determination of precision and bias in sediment testing of aquatic organisms is difficult or may be impossible in some cases, as compared to analytical (chemical) determinations. This is due, in part, to the many unknown variables which affect organism response. For a detailed discussion of precision as it relates to sediment toxicity testing, see Section 17 in Test Method E 1706.

14.1.1 *Bias*—The bias of toxicity tests cannot be determined since there is no acceptable reference material. The bias of the reference-toxicity tests can only be evaluated by comparing test responses to control charts be evaluated by comparing test responses to control charts. For a detailed discussion of bias as it relates to sediment toxicity testing, see Section 17 in Test Method E 1706.

14.1.2 The sensitivity of a toxicity test will depend upon the number of replicates per concentration or treatment, the variability within that treatment (among replicates), the probability levels (alpha and beta) and the statistical test used. Tests with anuran larvae have demonstrated that variability may occur within a treatment. This is especially the case for sublethal growth parameters where particularly small or large organisms can occur within a single treatment. Such differences in size may represent natural physiological differences (that is, poor health) or behavioral differences in individuals that affect access and consumption to available food and subsequent lower growth rates. The presence of unusually small or large specimens within a replicate chamber is to be occasionally expected and is not reason to discard individual measurements as outliers, unless all or most individuals in a single replicate exhibit mortality or growth patterns that are substantially different from other replicates within a treatment. Such a situation may indicate poorly homogenized sediment, technician error at test initiation or the presence of a highly-consolidated particle containing a toxic substance that is not representative of the sediment as a whole. In such cases, an outlier test may be appropriate to determine whether the replicate should be excluded from analysis. Exclusion of replicates should be avoided, however and every effort should be made collect enough sediment for a full eight replicates, in order to increase the statistical power of the test and reduce the effects of replicate variability **(50)**.

14.1.3 Intralaboratory precision data are routinely calculated for test organisms using water-only exposures to a reference toxicant, such as NaCl or KCl (as described in 11.6). Intralaboratory precision data should be tracked using a control chart. For reference toxicant tests with anurans, both survival and growth parameters should be tracked. Reference toxicant tests should be of a sufficient duration to achieve measurable growth (relative to the size of organisms at test initiation). For anurans, a minimum of seven days is recommended. Each

laboratory's reference-toxicant data will reflect conditions unique to that facility, including dilution water, culturing, and other variables. The conditions for the reference toxicant test, such as water type, test containers, organism age, feeding and concentration series, should remain the same. Altering test variables will introduce variation, wider confidence intervals and will compromise the integrity and usability of the reference toxicant data as a means of tracking intralaboratory precision.

14.1.4 Before conducting tests with potentially contaminated sediment, it is strongly recommended that the laboratory conduct the tests with control sediment(s) alone. Results of these preliminary studies should be used to determine if the use of the control sediment and other test conditions (that is, water quality) result in acceptable performance in the tests. If organism performance in the selected control sediment is inconsistent, an alternative sediment should be selected.

15. Keywords

15.1 amphibian; bioavailability; *Bufo* spp.; hydric soils; *Rana* spp.; *Rana pipiens*; sediment; toxicity; wetland

APPENDIXES

(Nonmandatory Information)

X1. LIST OF ALTERNATIVE SPECIES

X1.1 *Use of Alternative Species*—Although this procedure was developed with *R. pipiens*, it might be necessary to use alternative species when required by regulation or limited by seasonal availability of test organisms. Deviations from the procedures outlined in Table 4 should be recorded and it may be difficult to compare data between toxicity tests conducted with *R. pipiens* and alternative species.

X1.2 *Recommended Anurans*—Other members of the family Ranidae (for example, *R. sphenocephala*, *R. palustris*, or *R. catesbiena*) and Bufonidae (for example, *B. americanus* or *B. fowleri*) might be best suited for conducting a whole-sediment exposure toxicity test due to the commercial availability of eggs. High egg production, relevant geographical range, short hatching periods, and sensitivity to contaminants should be considered in selecting alternative species. *Xenopus laevis* may be considered as an alternative species due to the generally consistent availability of eggs; however, researchers should review existing data on the relative sensitivity to some contaminants **(49)**.

X1.2.1 Standard E 1439-98 includes a methodology for exposing *X. laevis* to whole sediments (referred to as solid phase sample testing). This methodology is an alternative to FETAX studies conducted in aqueous solutions. Although *Xenopus* is not native to the United States, the standardized, FETAX testing protocol, the availability of test organisms, and ease of use of *Xenopus* in the laboratory has made it a popular test species for amphibian toxicity testing.

X1.2.1.1 The FETAX solid phase testing may be performed in 250 mL specimen bottles or similar capped vessels equipped with a 55 mL glass tube with Teflon mesh insert as the exposure chamber. For screening tests, 35 g of sediment (dry weight) should be placed in the bottom of the vessel, with the Teflon mesh insert added, and should be filled with 140 mL of FETAX Solution. Blastulae stage embryos are placed directly on the mesh insert that rests directly over the top of the soil or sediment in the sediment/water interface region. Four to six dilutions ranging from 0 to 100 % soil sample and a FETAX Solution control are typically tested. Each sample should be tested in triplicate. Solutions and soils or sediments should be changed every 24 hours of the four-day test. At the end of the four-day exposure period, surviving embryos should be preserved in 3 % (w/v) formalin (pH 7.0) and morphological characteristics evaluated using a dissecting scope. Growth may be determined using a digitizing software package.

X1.2.1.2 While the alternative FETAX methodology exposes young amphibians to sediments there are several differences relative to the test conditions presented in Table 4. Primarily these differences are related to test duration and the age of the test organisms. The FETAX test is a rapid test designed to identify developmental toxicants. It is conducted over a relatively short duration (4 d) with recently fertilized embryos (mid blastula to early gastrula) and evaluates malformations, in addition to mortality and growth. The test conditions presented in Table 4 indicate a longer test duration (10 d) with older test organisms (≤72 hours old). This methodology evaluates survival and growth of tadpoles exposed directly to sediment and overlying water. The FETAX methodology is conducted with an amphibian species that is not native to North America. Although *X. laevis* may be available with less seasonal variability, in some cases it may be preferable to conduct a toxicity test with a species that is native to the test site.

X2. DATA ANALYSIS

X2.1 *General*—Test Method E 1706 provides guidance on data analysis. The following sections briefly summarize this guidance. Mortality or apparent size reduction in any sediment treatment is not necessarily an indication of toxicity. Statistical analysis is used to determine if apparent differences are significant (**52-54**). Organism response to test sediments is typically compared to the control response. If a reference sediment (for example, upstream or independent of a study site) is also collected, then test sediment results may be statistically compared against the reference sediment. Two types of data are obtained from the toxicity test: acute (mortality) and chronic (width and length). Each data type should be analyzed independently. If other measurements are also obtained (for example, weight or tissue burden) then those data can also be analyzed separately.

X2.2 *Forms of Evaluation*—Data analysis is in two general forms: hypothesis testing and point estimation. Hypothesis testing involves assigning an alpha level for the analysis and then, using that criterion, determining which treatments are significantly different from the control. If only field-collected sediment is tested, then data analysis will typically consist only of hypothesis testing. If however, a series of sediment dilutions were prepared (that is, mixing test sediment with control sediment at fixed percentages [6.25, 12.5, 25, 50]), or if spiked-sediment samples are prepared representing a true concentration gradient for chemical(s) of concern, then point estimates can be made. A point estimate, such as an LC50, is a concentration of test media at which a certain effect (for example, half the test organisms die) is determined to occur. General guidance for conducting these analyses is given in the following sections.

X2.2.1 *Hypothesis Testing*—Hypothesis testing should follow the same general structure as described by Test Method E 1706 and by U.S. Environmental Protection Agency (**1, 34**). In summary, mortality/survival data are analyzed first. If there is a significant reduction in survival in any treatment, that treatment is dropped from analysis of sublethal data. Determination of significant effects is dependent upon the predetermined alpha level. The alpha level, or α, is defined as the probability of committing a Type I statistical error—rejecting the null hypothesis (H_o) of no effect, even if H_o is true. That is, concluding a sample is toxic, even when it isn't (Table X2.1).

X2.2.1.1 The majority of studies in environmental toxicology are analyzed with an α of 0.05, which means there is a theoretical 5 % chance that a Type I error will be committed. The α level is not fixed and can be changed, depending upon the objectives of the study. A lower α—0.01 for example—will reduce the likelihood of a Type I error. However, it will also increase the likelihood of a Type II error (β), that is, concluding that a sample is not toxic when it, in fact, is. Historically, β and its inverse (1-β), which is the associated power of the test, have generally been ignored by environmental researchers. However, because the power of a test is defined as the probability of correctly detecting a true toxic effect, considering β may be important in designing a study. If α is held constant, for example, β decreases (and test power increases) as the sample size increases and variance decreases (**50**).

TABLE X2.1 Statistical Errors

Decision	If H_o is True	If H_o is False
H_o Rejected	Type I error (α)	No error
H_o Accepted	No error	Type II error (β)

X2.2.1.2 Since survival data often demonstrate non-normal distributions, proportional survival data are first transformed using an arc sine-squareroot transformation. The normality and homogeneity of variance are then evaluated using tests such as Shapiro-Wilk's and Bartlett's, respectively. If data are found to meet the normality and homogeneity of variance requirements of parametric tests, then differences from the control can be analyzed with Dunnett's Procedure (for an equal number of replicates) or a T-Test with Bonferroni adjustments (for unequal replicates). If data do not meet the assumptions for a parametric test, then nonparametric (rank) tests have to be used. The most common tests are Steel's Many-One Rank Test (for equal replicates) or Wilcoxon Rank Sum Test with Bonferroni adjustments (for unequal replicates).

X2.2.1.3 While these statistical tests are the ones most commonly used in the analysis of toxicity data, they are not the only ones available. For example, the objective of the study may be to determine if test sediments are significantly different from each other, as well as from the control. In that case, analysis of variance with Tukey's multiple range test (parametric) or a Kruskal-Wallis test (nonparametric) may be appropriate. Because of the many tests that are available, it is important that the project goals be thoroughly defined before data are collected.

X2.2.1.4 Sublethal effects are analyzed after mortality effects have been evaluated. Individual sublethal measurements are averaged to produce a mean width and length (per surviving organism) for each replicate. If there was significant mortality in any test treatment, that treatment is typically dropped from analysis of sublethal effects. Sublethal measurements are continuous data and therefore do not need to be transformed (arc sine-squareroot) before analysis. With that exception, the analysis of sublethal endpoints is the same as for survival.

X2.2.2 *Point Estimates*—Point estimations for individual chemicals of concern are seldom used in sediment tests conducted with field-collected samples because there is generally not a single concentration gradient for the particular chemical of concern. In addition, field-collected sediments may contain multiple toxicants that could act independently or have synergistic, additive, or antagonistic effects. For example, if a sediment (for example, from a historical mining district) has high concentrations of copper, zinc, and cadmium, all of which may be at toxic levels, a point estimate based on the concentration of any one metal may be meaningless because of the presence of the other metals. However, point estimates could

be calculated based upon the percent (weight or volume) of a test sediment mixed with a nontoxic control sediment. If this method is used, then both sediments should have about the same moisture fraction so that the percentage estimates are reasonably accurate. Point estimates could also be used if samples are collected along a known concentration gradient for one particular chemical and no other chemicals of concern are present. Finally, if spiked sediment tests are conducted where different treatments of sediment contain variable but known quantities of a particular chemical, then point estimates can be made.

X2.2.2.1 Any of the point estimation procedures calculate a concentration (mass per volume or percent) at which a certain effect will occur. An LC50, for example, is the concentration at which 50 % of the organisms are expected to die while an IC25 is the concentration which causes a 25 % reduction in the endpoint of interest. The manner in which LC50s or other point estimates are calculated varies with the structure of the data. For example, if the responses in the test treatments are all or nothing (either everything is alive or everything is dead), than the simplest method—graphical—is used. LC50s using the graphical method, like the name implies, are calculated on graph paper, although a simpler method is simply calculating the geometric mean of the highest "all-alive" concentration and the lowest "all-dead" concentration. If there is partial mortality in any test treatment then a Spearman-Karber, Trimmed Spearman-Karber, or Probit method should be used. These methods are described in detail by U.S. Environmental Protection Agency (**55**). In brief, if there are two or more treatments with partial mortality, then use of the Probit method (parametric) is indicated. In situations where the Probit method is inappropriate due to non-normal or significantly heterogeneous data, the Trimmed Spearman-Karber or Spearman-Karber Methods may be used. These LC50 procedures are available with a variety of computer software programs (**52-54**).

X2.2.2.2 LC50 models, by definition, are used to calculate point estimates for mortality endpoints, although the models can also be used to calculate point estimates for nonlethal endpoints (for example, median effects concentrations (EC50s)). The Linear Interpolation Method was developed for the general application to data generated during chronic toxicity tests. The endpoint generated by the Linear Interpolation Method is an ICp value, where IC = Inhibition Concentration and p is the percent effect. The value of p can be adjusted, although the most typical values are 25 and 50. The Linear Interpolation Model assumes a linear response from one concentration to the next and assumes that the mean response of the next higher concentration will be equal to or less than the preceding concentration. If this is not the case, the data are adjusted by smoothing. A more thorough discussion of the Linear Interpolation Model is provided by Norberg-King (**56**).

REFERENCES

(**1**) U.S. Environmental Protection Agency, "Methods for Measuring the Toxicity and Bioaccumulation of Sediment-Associated Contaminants with Freshwater Invertebrates–Second Edition," EPA/600/R-99/064, 2000.

(**2**) U.S. Environmental Protection Agency–U.S. Army Corps of Engineers, "Evaluation of Dredged Material Proposed for Ocean Disposal–Testing Manual," EPA 503/8-91/001, 1991.

(**3**) U.S. Environmental Protection Agency–U.S. Army Corps of Engineers, "Evaluation of Dredged Material Proposed for Discharge in Waters of the U.S.–Testing Manual," EPA 823/B-98/004, 1998.

(**4**) Francis, P. C., Birge, W. J., and Black, J. A., "Effects of Cadmium-Enriched Sediment on Fish and Amphibian Embryo-Larval Stages," *Ecotoxicology and Environmental Safety*, Vol 8, 1984, pp. 378–387.

(**5**) Berrill, M., Coulson, D., McGillivray, L., and Pauli, B., "Toxicity of Endosulfan to Aquatic Species of Anuran Amphibians," *Environmental Toxicology and Chemistry*, Vol 17, 1998, pp. 1738–1744.

(**6**) Edginton, A. N., Stephenson, G. R., Sheridan, P. M., Thompson, D. G., and Boermans, H. J., "Effect of pH and Release® on Two Life Stages of Four Anuran Amphibians," *Environmental Toxicology and Chemistry*, Vol 22, 2003, pp. 2673–2678.

(**7**) Jofre, M. B. and Karasov, W. H., "Direct Effect of Ammonia on Three Species of North American Anuran Amphibians," *Environmental Toxicology and Chemistry*, Vol 18, 1999, pp. 1806–1812.

(**8**) Snodgrass, J. W., Hopkins, W. A., Jackson, B. P., Baionno, J. A., and Broughton, J., "Influence of Larval Period on Responses of Overwintering Green Frog (*Rana clamitans*) Larvae Exposed to Contaminated Sediments," *Environmental Toxicology and Chemistry*, Vol 24, 2005, pp. 1508–1514.

(**9**) Leney, J. L., Drouillard, K. G., and Haffner, G. D., "Does Metamorphosis Increase the Susceptibility of Frogs to Highly Hydrophobic Contaminants?" *Environmental Science and Technology*, Vol 40, 2006, pp. 1491–1496.

(**10**) Ankley, G. T., Tietge, J. E., DeFoe, D. L., Jensen, K. M., Holcombe, G. W., Durhan, E. J., and Diamond, S. A., "Effects of Ultraviolet Light and Methoprene on Survival and Development of *Rana pipiens*," *Environmental Toxicology and Chemistry*, Vol 17, 1998, pp. 2530–2542.

(**11**) James, S. M. and Little, E. E., "The Effects of Chronic Cadmium Exposure on American Toad (*Bufo americanus*) Tadpoles," *Environmental Toxicology and Chemistry*, Vol 22, 2003, pp. 377–380.

(**12**) U.S. Environmental Protection Agency, "Method for Assessing the Chronic Toxicity of Marine and Estuarine Sediment-Associated Contaminants with the Amphipod *Leptocheirus plumulosus*," EPA/600/R-01/020, 2001.

(**13**) Naval Facilities Engineering Command (NAVFAC), "Development of a Standardized Approach for Assessing Potential Risks to Amphibians Exposed to Sediments and Hydric Soils," March, 2004.

(**14**) Naval Facilities Engineering Command (NAVFAC), "Laboratory Validation Plan for ESTCP Project CU-0514 Demonstration and Certification of Amphibian Ecological Risk Assessment Protocol," December 2005.

(**15**) Loveridge, A. R., Lavoie, D. R., Corl, W. E., Jackson, R., Speicher, J., Pillard, D. A., and Christopher, J., "Use of a New Standardized Test for Assessing Potential Risks to Amphibians Exposed to Whole Sediments," Poster presented at Annual Meeting of Society of Environmental Toxicology and Chemistry, Portland, OR. November 2004.

(**16**) Environmental Chemical Corporation, "Draft Ecological Risk Assessment Addendum–Area of Concern CS-10 Detail F, Massachusetts Military Reservations, Cape Cod, MA," October, 2004.

(**17**) ENSR, "RAM Completion Statement & RAO Class A-2 Statement, Wetlands Reclamation Project Adjacent to MassHighway Maintenance Depot, Marrett Road (Rte. 2A) Lexington, MA MADEP RTN: 3–20801," November, 2002.

(18) ENSR, "Draft Baseline Ecological Risk Assessment, Sixty-One Industrial Park Site Memphis, Tennessee," May, 2007.

(19) Lande, S. P. and Guttman, S. I., "The Effects of Copper Sulfate on the Growth and Mortality Rate of *Rana pipiens* Tadpoles," *Herpetologia*, Vol 29, 1973, pp. 22–27.

(20) Birge W. A. and Black, J. A., "Effects of Copper on Embryonic and Juvenile Stages of Aquatic Animals, " J. O. Nriago (ed.), *Copper in the Environment*, 1979, pp. 373–399.

(21) Khangarot, B. S., Sehgal, A., and Bhasin, M. K., "Man and Biosphere–Studies on the Sikkim Himalayas. Part 5: Acute Toxicity of Selected Heavy Metals on the Tadpoles of *Rana hexadactyla*," *Acta Hydrochimica et Hydrobiologica*, Vol 13, 1985, pp. 259–263.

(22) Mudgall, C. F. and Patil, S. S., "Toxicity of Lead and Mercury to Frogs *Rana cyamophlyotis* and *Rana tigerina*," *Environment and Ecology*, Vol 6, 1988, pp. 506–507.

(23) Taylor, D. H, Steele, C. W., and Strickler-Shaw, S., "Responses of Green Frog (*Rana clamitans*) Tadpoles to Lead-Polluted Water, " *Environmental Toxicology and Chemistry*, Vol 9, 1990, pp. 87–93.

(24) Horne, M. T. and Dunson, W. A., "Effects of Low pH, Metals and Water Hardness on Larval Amphibians," *Arch. Enviro. Contam. Toxicol*, Vol 29, 1995, pp. 500–505.

(25) Fioramonti, E., Semlitsch, R. D., Reyer, Heinz-Ulrich, and Fent, K., "Effects of Triphenyltin and pH on the Growth and Development of *Rana lessonae* and *Rana esculenta* Tadpoles," *Environmental Toxicology and Chemistry*, Vol 16, 1997, pp. 1940–1947.

(26) Hall Jr., L. W., Scott, M. C., and Killen, W. D., "Ecological Risk Assessment of Copper and Cadmium in Surface Waters of Chesapeake Bay Watershed," *Environmental Toxicology and Chemistry*, Vol 17, 1998, pp. 1172–1189.

(27) Howe, C. M., Berrill, M., Pauli, B. D., Helbing, C., Werry, K., and Veldhoen, N., "Toxicity of Glyphosate-based Pesticides to Four North American Frog Species," *Environmental Toxicology and Chemistry*, Vol 23, 2004, pp. 1928–1938.

(28) Jung, R. E. and Walker, M. K., "Effects of 2,3,7,8-Tetrachlorodibenzo-*p*-Dioxin (TCDD) on Development of Anuran Amphibians," *Environmental Toxicology and Chemistry*, Vol 16, 1997, pp. 230–240.

(29) Allran, J. W. and Karasov, W. H., "Effects of Atrazine and Nitrate on Northern Leopard Frog (*Rana pipiens*) Larvae Exposed in the Laboratory from Posthatch through Metamorphosis," *Environmental Toxicology and Chemistry*, Vol 19, 2000, pp. 2850–2855.

(30) Wojtaszek, B. F., Buscarini, T. M., Chartrand, D. T., Stephenson, G. R., and Thompson, D. G., "Effect of Release® Herbicide on Mortality, Avoidance Response, and Growth of Amphibian Larvae in Two Forest Wetlands," *Environmental Toxicology and Chemistry*, Vol 24, 2005, pp. 2533–2544.

(31) Fort, D. J., Propst, T. L., Stover, E. L., Helgen, J. C., Levey, B., Gallagher, K., and Burkhart, J. G., "Effects of Pond Water, Sediment, and Sediment Extracts from Minnesota and Vermont, USA, on Early Development and Metamorphosis of *Xenopus*," *Environmental Toxicology and Chemistry*, Vol 18, 1999, pp. 2305–2315.

(32) Savage, W. K., Quimby, F. W., and DeCaprio, A. P., "Lethal and Sublethal Effects of Polychlorinated Biphenyls on *Rana sylvatica* Tadpoles," *Environmental Toxicology and Chemistry*, Vol 21, 2002, pp. 168–174.

(33) Widder, P. D. and Bidwell, J. R., "Cholinesterase Activity and Behavior in Chlorpyrifos-exposed *Rana sphenocephala* Tadpoles," *Environmental Toxicology and Chemistry*, Vol 25, 2006, pp. 2446–2454.

(34) U.S. Environmental Protection Agency, "Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms-Fourth Edition," EPA-821-R-02-013, 2002.

(35) U.S. Environmental Protection Agency, "National Recommended Water Quality Criteria," United States Environmental Protection Agency, 2006, p. 24.

(36) Chapman, P. M., "Current Approaches to Developing Sediment Quality Criteria," *Environmental Toxicology and Chemistry*, Vol 8, 1989, pp. 589–599.

(37) Di Toro, D. M., Zarba, C. S., Hansen, D. J., Berry, W. J., Swartz, R. C., Cowan, C. E., Pavlou, S. P., Allen, H. E., Thomas, N. A., and Paquin, P. R., "Technical Basis for Establishing Sediment Quality Criteria for Nonionic Organic Chemicals Using Equilibrium Partitioning," *Environmental Toxicology and Chemistry*, Vol 10, 1991, pp. 1541–1583.

(38) Maloney, J., "Influence of Organic Enrichment on the Partitioning and Bioavailability of Cadmium in a Microcosm Study," *Marine Ecology Progress Series*, Vol 144, 1996, pp. 147–161.

(39) Granberg, M. E. and Forbes, T. L., "Role of Sediment Organic Matter Quality and Feeding History in Dietary Absorption and Accumulation of Pyrene in the Mud Snail (*Hydrobia ulvae*)," *Environmental Toxicology and Chemistry*, Vol 25, 2006, pp. 995–1006.

(40) Ankley, G. T., Benoit, D. A., Balogh, J. C., Reynoldsen, T. B., Day, K. E., and Hoke, R. A., "Evaluation of Potential Confounding Factors in Sediment Toxicity Tests with Three Freshwater Benthic Invertebrates," *Environmental Toxicology and Chemistry*, Vol 13, 1994, pp. 627–635.

(41) Sibley, P. K., Ankley, G. T., Cotter, A. M. and Leonardi, E. N., "Predicting Chronic Toxicity of Sediments Spiked with Zinc: An Evaluation of the Acid Volatile Sulfide (AVS) Model Using a Lifecycle Test with the Midge (*Chironomus tentans*)," *Environmental Toxicology and Chemistry*, Vol 15, 1996, pp. 2102–2112.

(42) Ditsworth, G. R. and Schults, D. W., "Preparation of Benthic Substrates for Sediment Toxicity Testing," *Environmental Toxicology and Chemistry*, Vol 9, 1990, pp. 1523–1529.

(43) Hourdry, J., L'Hermite, A., and Ferrand, R., "Changes in the Digestive Tract and Feeding Behavior of Anuran Amphibians During Metamorphosis," *Physiological Zoology*, Vol 69, 1996, pp. 219–251.

(44) Gosner, K. L., "A Simplified Table for Staging Anuran Embryos and Larvae with Notes on Identification," *Herpetologica*, Vol 16, 1960, pp. 183–190.

(45) Stemmer, B. L., Burton Jr., G. A., and Sasson-Bricksen, G., "Effect of Sediment Spatial Variance and Collection Method on Cladoceran Toxicity and Indigenous Microbial Activity Determinations," *Environmental Toxicology and Chemistry*, Vol 9, 1990, pp. 1035–1044.

(46) Sarda, N. and Burton Jr., G. A., "Ammonia Variation in Sediments: Spatial, Temporal and Method-Related Effects," *Environmental Toxicology and Chemistry*, Vol 14, 1995, pp. 1499–1506.

(47) Landrum, P. F., Eadie, B. J., and Faust, W. R., "Variation in the Bioavailability of Polycyclic Aromatic Hydrocarbons to the Amphipod Diporeia with Sediment Aging," *Environmental Toxicology and Chemistry*, Vol 11, 1992, pp. 1197–1208.

(48) Karasov, W. H., Jung, R. E., Langenberg, S. V., and Bergeson, T. L. E., "Field Exposure of Frog Embryos and Tadpoles Along a Pollution Gradient in the Fox River and Green Bay Ecosystem in Wisconsin, USA," *Environmental Toxicology and Chemistry*, Vol 24, 2005, pp. 942–953.

(49) Birge, W. J., Westerman, A. G., and Spromberg, J. A., "Comparative Toxicology and Risk Assessment of Amphibians," *Ecotoxicology of Amphibians and Reptiles*, Sparling, Linder and Bishop, eds., Society of Environmental Toxicology and Chemistry, Pensacola, FL, 2000, pp. 727–791.

(50) Denton, D. L. and Norberg-King, T .J., "Whole Effluent Toxicity Statistics: A Regulatory Perspective," *Whole Effluent Toxicity Testing: An Evaluation of Methods and Prediction of Receiving System Impacts*, Grothe, Dickson, and Reed-Judkins, eds., Society of Environmental Toxicology and Chemistry, Pensacola, FL, 1996, pp. 83–102.

(51) Shumway, W., "Stages in the Normal Development of Rana pipiens," *The Anatomical Record*, Vol 78, 1940, pp. 139–148.

(52) WEST, Inc. and D. D. Gulley, "Toxstat Version 3.5," Western Ecosystems Technology, Inc. Cheyenne, WY, 1996.

(**53**) Analytical Software, "Statistix 7 for Windows," Analytical Software, Tallahassee, FL, 2000.
(**54**) U.S. Environmental Protection Agency, "U.S. EPA Toxicity Data Analysis Software. Version 1.5," Environmental Monitoring Systems Laboratory, Cincinnati, OH, 1994.
(**55**) U.S. Environmental Protection Agency, "Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms–Fifth Edition," EPA-821-R-02-012, 2002.
(**56**) Norberg-King, T., "A Linear Interpolation Method for Sublethal Toxicity: The Inhibition Concentration (ICp) Approach (Version 2.0)," National Effluent Toxicity Assessment Center Technical Report 03-93, U.S. Environmental Protection Agency, Duluth, MN, 1993.