INTEGRATED FIELD-SCREENING FOR RAPID SEDIMENT CHARACTERIZATION

ESTCP PROJECT #9717

Space and Naval Warfare Systems Center, San Diego

September 30, 2000

Environmental Security
Technology Certification
Program

Table of Contents

List of Tables

List of Figures

1. Introduction

1.1 Background Information.

The Environmental Security Technology Certification Program (ESTCP) has an established program to accelerate acceptance and application of innovative monitoring and site characterization technologies that improve the way the nation manages its environmental problems. Space and Naval Warfare Systems Center, San Diego (SSC San Diego), will demonstrate an integrated methodology to facilitate acceptance of three field screening techniques to delineate chemical concentrations and potential biological effects of sediment contaminants.

Defining the nature and extent of contamination in marine sediments can be a difficult problem. Detailed site investigations require extensive sampling and subsequent laboratory analyses for both metal and organic contaminant chemistries. Additional laboratory analyses including several different types of bioassays are conducted to determine any possible adverse biological effects to organisms exposed to the sediment. Samples are often collected without any *a priori* knowledge of the nature and extent of contamination. Due to the high cost of all these laboratory analyses, samples taken for analysis are often limited. Zones of contamination in marine sediments can be missed, or, if located, over- or under-estimated. For more detailed spatial information on extent of contamination, sites of interest must often be sampled and analyzed in an iterative manner. This approach can be prohibitively costly, slow and labor-intensive.

An alternative to this approach is to combine standard laboratory analyses with field screening using a number of techniques to characterize both the contaminant chemistry and any possible biological effects. By utilizing near real-time screening techniques during the sampling procedure, the full extent of contamination and any possible biological effects can be rapidly mapped. This also allows more informed selection of a subset of the screened samples to continue on for laboratory analyses to fully describe the nature of contamination and biological effects. The use of geo-statistical procedures can provide the basis for selecting sampling strategies and aid in selecting the number of samples to be screened and which samples will also continue onto laboratory analysis (Chapter 9 (“Double Sampling”) in Gilbert, 1987).

Field Portable X-ray Fluorescence (FPXRF) Spectrometry will be used to screen for metals of interest. This technique measures the fluorescence spectrum of x-rays that occur when atoms are excited by a radiation source. The energy of the emitted x-rays reveal the identity of the metals in the sample, and the intensity of emitted x-rays is related to their concentrations (Swift, 1995). Rapid, multi-element analysis can be performed by XRF. An XRF spectrometer can analyze a wide range of elements (i.e., sulfur through uranium), with a wide dynamic range, from parts per million to percent levels, encompassing typical element levels found in soils and sediments. FPXRF units provide near real time measurements with minimal sample handling, allowing for extensive, semi-quantitative analysis on site. Several examples can be found in literature in which FPXRF has been used for the analysis of soils and sediments. Sediments in a Norwegian fjord (Skei *et al*., 1972), San Diego Bay (Stallard *et al.*, 1995), and a large number of sites (Kirtay *et al*., 1997, 1998) have been screened for heavy metals content by XRF. FPXRF has

been certified by the USEPA as a field screening method for metals in soils (EPA SW-846 Method 6200).

Polycyclic aromatic hydrocarbons (PAHs) are common organic contaminants and in addition to metals, are found in many contaminated sediments. PAHs in the sediments may originate from many sources due to incomplete combustion of wood or petroleum products and from petroleum itself. They are common constituents in automobile tires (and therefore stormwater) and creosote (and therefore pier pilings). To screen for PAHs of interest, UV Fluorescence (UVF) Spectroscopy will be used to estimate total PAH levels in the sediments. The screening method is based on measurement of the amount of fluorescence observed following UV excitation of solvent extracts separated from bulk sediments. Because many different PAHs can fluoresce, this method is not able to quantify individual PAHs, but can serve to screen for bulk PAH levels in the sediment. Many studies have used UVF to assess total PAH levels in various types of sediment (Hargrave and Phillips 1975; Filkins 1992; Owen *et al*. 1995).

Although contaminants may be present in the sediment, bulk chemistry measurements cannot determine if these contaminants are bioavailable and therefore hazardous to organisms. There is a general requirement to use specified toxicity tests, such as the 10-day solid phase amphipod test or 48 hour pore water sea urchin development test, to meet regulatory needs. These tests are lengthy and expensive. To screen for biological effects, a technique based on the QwikSed screening bioassay will be employed (Lapota *et al*., 1997). The QwikSed test on pore water samples shows good correlation to the sea urchin development test from preliminary samples, and has the added advantage of time and cost savings. The QwikSed Bioassay measures the inhibition of light emitted by the marine bioluminescent dinoflagellate, *Gonyaulax polyedra*, exposed to a test solution (effluents, elutriates, or sediment pore waters). Any decrease in light output relative to controls is suggestive of the presence of bioavailable contaminants.

This final report describes how the demonstration team collected and analyzed samples to verify the field screening technologies and ensure that demonstration activities would be documented and scientifically sound, and that performance data of known quality would be collected. SSC San Diego has prepared this technology demonstration plan following the guidelines in the ESTCP Program Offices' document; “Final Report Guidelines for Funded Projects" dated April 15, 1996. The final report is divided into 10 sections. Section 1 provides a broad overview of the purpose and background of the demonstration and a description of the technology demonstration process. Section 2 provides technology descriptions and background on the three screening technologies. Section 3 provides a description of the demonstration sites. Section 4 presents the demonstration approach. Section 5 presents an assessment of the performance of the screening techniques with discussions of the demonstration data along with standard laboratory validation data. Section 6 presents the cost performance data to allow estimates of actual costs to use screening techniques. Section 7 presents the regulatory issues. Section 8 presents some thoughts on technology implementation and the efforts being used to get these screening technologies into standard regulatory projects. Section 9 lists some of the lessons learned and Section 10 lists the references.

1.2 Official DoD Requirement Statement.

1.3 Objectives of the Demonstration.

The three field screening technologies were demonstrated at two demonstration sites to facilitate their acceptance by regulatory agencies. The first demonstration was at the Naval Air Station Alameda located just outside of Oakland, California. A second demonstration was held at two sites within the Naval Complex Pearl Harbor in Hawaii.

The purpose of the demonstrations was to generate field data appropriate for verifying the integrated performance of the three field screening technologies, and thereby facilitate the technologies' acceptance and use by the regulator and user communities so that field screening becomes integrated into standard assessment of sediments. As part of the demonstrations we collected approximately 50 to 100 samples (depending on the particular screening technique) to be used in validation of the screening techniques for the ESTCP project. Additionally, we have collected a similar number of pre-demonstration samples that may also be used for validation.

The primary objectives of this demonstration are to evaluate the three field screening technologies in the following areas: (1) their performance compared to conventional sampling and analytical methods; (2) data quality; (3) the logistical and economic resources necessary to operate the technologies; and (4) the range of usefulness in which the technologies can be operated and integrated into a screening procedure that allows more efficient assessment of sediment sites. Secondary objectives for this demonstration are to evaluate the technologies for their reliability, ruggedness, and ease of operation. The fourth primary objective is important because current regulatory projects often rely on “blind” sampling, with little or no knowledge of how much volume of sediment each laboratory measurement represents. Cheaper screening techniques will allow more knowledgeable sample selection for laboratory analysis, and therefore, better insight into how representative these samples are. As additional screening techniques are developed, they may be incorporated into existing screening procedures.

Performance of field screening technologies is often evaluated to determine the percentage agreement between "detect/non-detect" field screening data and corresponding laboratory results. Because of natural interferences and little or no sample preparation, a greater number of false positives and false negatives are expected with field screening technologies when compared to laboratory techniques. The three screening techniques discussed here are able to provide more information than simple “detect/non-detect” data, and are more accurately termed “semi-quantitative” techniques. Therefore, in addition to reporting % detected, % non-detected, % false positives, and % false negatives, we will also be using performance criteria based on correlation coefficient between field screening and laboratory data (See Section 5).

1.4 Regulatory Issues.

The assessment of contamination in sediment requires both contaminant chemistry and biological effects information. Laboratory chemistry analysis and bioassays are typically used to generate the required data. This project will demonstrate how field screening data can provide the regulatory community with supplemental information on the extent of contamination with more extensive and higher density data that is obtained both faster and cheaper than current practices normally allow. The detection levels of the screening techniques (see Section 2.1)

provide data relevant to the regulatory criteria that are often used at sediment sites. For example, adverse biological effects to organisms begin to be observed when total PAH values are above 4 ppm in the sediments (Long *et al.*, 1995). Detection levels for PAHs by screening with UV fluorescence techniques are also in this range (Owen *et al.* 1995, our data in Section 5).

For the metals copper, zinc, and lead, total individual metal levels are the regulatory criteria against which the screening technique will be evaluated. Because not all of the total levels of contaminants are bioavailable, the screening bioassay can be used to infer what fraction of total contaminant is actually bioavailable. The screening bioassay will be evaluated against standard laboratory bioassay endpoints such as percent survival.

1.5 Previous Testing of the Technology.

An ongoing validation process of the three field screening technologies has been used by SSC San Diego at multiple sites to provide a database for review by regulatory agencies in technology acceptance programs. The XRF, UVF, and QwikSed techniques are currently being either evaluated or demonstrated independently for different matrices by several technology certification programs including the following:

FPXRF - California Environmental Protection Agency (Cal-EPA) - Technology Certification Program

UVF - USEPA, Department of Defense, and Department of Energy - Consortium for Site Characterization Technology

QwikSed - American Society for Standard Tests and Methods, (ASTM)

2. Technology Description

2.1 Description.
This section describes the FPXRF, UVF, and QwikSed technologies including background information and description of equipment. General theories, functionality, and operations associated with the technologies are also discussed.

2.1.1 FPXRF.
XRF technologies operate on the concept of energy dispersive x-ray fluorescence spectrometry, a nondestructive qualitative and quantitative analytical technique. Most field portable XRF units use sealed radioisotope sources to irradiate samples with x-rays. Laboratory-grade XRF technologies generally use an x-ray tube to irradiate the samples with x-rays, and both the field portable and laboratory-grade technologies produce x-rays of known energies. By exposing a sample to an x-ray excitation source having energy close to, but greater than, the binding energy of the inner shell electrons of the metals, an inner shell electron is discharged. Electrons cascading in from outer electron shells fill the electron vacancies that result. Electrons in outer shells have higher energy states than inner shell electrons; therefore, to fill the vacancies, the outer shell electrons give off energy in the x-ray spectrum as they cascade down into the inner shell vacancies. There are three electron shells generally involved in the emission of x-rays during the XRF analysis of environmental samples: K, M, and L shell electrons. The emission of x-rays is termed x-ray fluorescence. Each metal gives off x-rays of specific energy levels. The specific type or energy of the emitted x-ray is unique to a given metal and is called a "characteristic" x-ray. By measuring the different energies of x-rays emitted by a sample exposed to an x-ray source, it is possible to identify and quantify the metals composition of a sample (Bertin 1975, Russ 1984).

An EPA Method (Method 6200) for the determination of metals in soils and sediments by Field-Portable X-ray Fluorescence Spectrometry (FPXRF) has been prepared by the USEPA (EPA, 1998). This method outlines the capabilities of the portable instrument relative to detailed standard chemical analyses and evaluates instrument response for sediments with different characteristics including moisture content, mineralogy, and modes and sources of contamination.

General Design Criteria:

- Source to provide x-rays
- Sample presentation device
- Detector that converts x-ray generated photons emitted from the sample into measurable electronic signals
- Data processing to include an emission or fluorescence energy analyzer

Analyses for this demonstration will be performed using a TN Spectrace 9000® portable XRF spectrometer (TN Spectrace Instruments). The instrument contains three radioisotope sources, Fe-55, Cd-109 and Am-241, to provide the excitation x-rays. It has an electronically cooled solid-state mercury iodide detector for measurement of the characteristic fluorescent x-rays. The instrument utilizes proprietary fundamental parameters (FP) algorithms, which eliminate the need for empirical calibration with site-specific standards.

One manufacturer-supplied FP application is used for sample analysis – "coarse-grain". The instrument is calibrated at the factory using pure elements, and computes contaminant concentrations using the matrix correcting FP algorithms. The coarse grain application is designed to correct for grain size effects and is used primarily for wet, heterogeneous sediment samples. Count times of 200-sec (Cd-109), 200 sec (Fe-55) and 20 sec (Am-241) are typically used for analyses, but these can vary depending on the amount of time available per analysis, the detection limits required and the age of the radioisotope sources.

This Spectrace 9000 FPXRF (Figure1) is capable of measuring up to 25 elements simultaneously. FPXRF measurements for Cu, Pb, and Zn will be the metals of concern in this demonstration. These heavy metals are common contaminants at military sites and are found in marine sediment at levels detectable by FPXRF. Other elements of concern (e.g., As, Cd, Cr, Hg, Ni) will not be considered because they are often present at concentrations not measurable by this XRF (*i.e*., below detection limits). Statistical methods were used in the laboratory to determine the detection limits of the instrument for Cu, Zn, and Pb (Kirtay *et al*., 1997). The Limit of Detection (LOD) was set at three times the within-batch standard deviation of replicate (n = 15) analyses of a blank sample (quartz). The LOD represents a criterion for a detection determination, i.e., deciding whether to classify a result as detected or not detected when the observed signal is close to that obtained for blank measurements. The manufacturer's stated TN Spectrace 9000 detection limits for Cu (44 ppm), Pb (14 ppm), and Zn (35 ppm) are lower than the ranges found in the laboratory testing with the coarse grain FP application. The ranges found here for marine sediments for Cu (50-100 ppm), Pb (25-50 ppm), and Zn (50-100 ppm) are comparable to what the USEPA found in their assessment of the TN Spectrace 9000 in soils (PRC, 1995).

Key Design Criteria of the Spectrace FPXRF:

- Radioisotope sources to produce x-rays
- Source detector beryllium window
- 1-inch diameter polypropylene probe window
- HgI_2 semi-conductor x-ray detector
- Multi-channel analyzer contained in electronics unit
- Hard-shell water repellant carrying case for protection

2.1.2 UVF.

Fluorescence can be used to measure the concentration of various organic analytes in addition to metals. Unlike metals, where high energy x-rays are required to generate fluorescence, PAHs require only UV light excitation to fluoresce visible light (Figure2). Excitation light from a lamp is passed through an excitation filter that transmits light of the chosen wavelength range. The light passes through the sample, causing the sample to emit light (fluoresce) proportional to the concentration of the fluorescent molecule (PAH) in the sample. The emitted light is passed through another optical filter (emission filter) before reaching the detector (in this case, a photomultiplier tube). The excitation wavelength is chosen (1) for strong absorption by the material under study, and (2) for minimal absorption by any interfering fluorescent materials that may be present. The photomultiplier and emission filters are also chosen so that (1) they respond as much as possible to the light emitted by the material under study, (2) they respond as little as

possible to the emission of any interfering fluorescent materials which may be present (see discussion in Turner Operating Manual, 1997).

General Design Criteria

- Source to provide UV excitation
- Light sensor / photomultiplier tube
- Sample chamber

PAHs will be screened in this study by UVF on solvent extracts of the sediments (Filkins 1992, Owen *et al*. 1995). A Turner fluorometer (Turner Model AU-10 Digital Filter Fluorometer®) with a standard optical package (commonly used in routine water quality analyses) will be used to screen for total PAHs (Figure 2). This optical package from the manufacturer is specifically designed for measurement of heavier weight PAH fluorescence with an excitation wavelength of 360 $\pm$ 10 nm provided by a quartz-halogen lamp. The detector system consists of a high gain, low noise photomultiplier tube with detection wavelength of 400-650 nm.

UV fluorescence is capable of screening for PAHs in a variety of matrices including water, tissue, soil, and sediment. Because fluorescence measurements are matrix sensitive, it is currently necessary to make measurements on solvent extracts rather than directly on the wet solid sediment sample. This solvent extraction step requires additional time for each sample analysis, so although fluorescence is a near real-time measurement the total time for analysis on solvent extracts may be up to half an hour. The solvent extraction step makes it possible to obtain an improvement in LODs of up to several orders in magnitude due to lower detection limits for PAHs in solution versus PAHs adsorbed on solid samples. LOD ranges from 1-5 ppm total PAH are based on comparisons to laboratory GC/MS data (see Section 5; Owen *et al.* 1995).

Key Design Criteria of the Turner UVF:

- Battery Power
- Multiple filter kits for various applications
- Internal Data Logging & Electronic Chart Recording
- Automatic range changing
- Rugged, water repellent field-portable instrument

2.1.3 QwikSed.

Many marine phytoplankton have the ability to produce bioluminescence, a visible blue light, as part of their daily physiological process. Traditional phytoplankton bioassays involve labor intensive enumerations of cells, however, a bioassay has been developed in our laboratory which makes use of the inherent bioluminescent characteristic of a particular group of phytoplankton called dinoflagellates (Lapota *et al*., 1994). Early observations indicated that the presence of some toxicants inhibited the amount of light produced by bioluminescent bacteria (Sie and Thanos 1977, Tchan and Chiou 1977, Johnson *et al*. 1942). More recently, bacterial bioluminescence toxicity test systems have been commercialized (Bulich 1979). The ecological role that dinoflagellates play (primary producers) makes them ideal for laboratory use, as they are unicellular, photosynthetic, and sensitive to toxicants. The QwikLite bioassay system was

recently developed to measure the light output from bioluminescent dinoflagellates for assessment of toxic effects when exposed to many chemicals, individually or in compounds, effluents, and antifoulant coatings. Successful bioassays of this type have provided data on acute response as well as chronic effects (from 3 hours up to 11 days) on two species of dinoflagellate, *Pyrocystis lunula* and *Gonyaulax polyedra*.

The basis of detection is to measure a light reduction from bioluminescent dinoflagellates following exposure to a toxicant. Bioluminescence is the production of light by living organisms due to an enzyme-catalyzed chemical reaction. Upon exposure to a toxicant, the dinoflagellates may shed an outer cell membrane called a theca and form a cyst. Consequently light production decreases from the dinoflagellates. Encystment is a normal response by dinoflagellates to an unfavorable or stressful environment.

The QwikLite or QwikSed (the sediment version instrument of the QwikLite) bioassay system consists of a horizontally-mounted 2 -inch diameter RCA 8575 photomultiplier tube (PMT) with an S-20 response used in the photon count mode. The QwikLite test chamber is constructed from black delron and is connected to the controller box via a combined power and signal cable. The top of the chamber is removable and houses a small adjustable stainless steel shaft terminating in a plastic propeller. The controller box has face displays for PMT and stirring motor voltages, PMT count LED, preset count time settings, manual and automatic switches to run the system, and backlit start, stop, and reset buttons. Neutral density optical filters can be easily changed (ND-1, ND-2, ND-3) to prevent PMT saturation. Dinoflagellate cells are cultured in optical grade spectrophotometric plastic cuvettes, which are placed individually into the test chamber (Figure 3).

General Design Criteria:

- Source to provide mechanical excitation
- Light sensor / photomultiplier tube
- Sample chamber

Cultures of dinoflagellates are maintained in a sterile enriched seawater medium (ESM) under 40-watt cool-white fluorescent bulbs on a 12:12 hr (light:dark) cycle at 19-20°C. Cells are cultured in ~600 mL ESM in borosilicate Erlenmeyer flasks at 2000-3000 cells/mL. Bioluminescent dinoflagellates are most stimulable and produce maximum light during the dark phase.

Sediment leachates are prepared by mixing sediments with filtered seawater in a 1:4 ratio for 1.5 hrs (USEPA, 1991). Pore waters can also be prepared by centrifugation. Both are diluted to 6.25% of the original sediment leachate or sediment pore water to find the endpoint or IC_{50} (concentration of the tested material which reduces or inhibits bioluminescence by 50% when compared to control cells). The salinity of the leachates and pore waters are checked and adjusted with sea salts to a standard salinity of 33 ppt prior to dispensing. Leachates and pore waters are diluted to 6.25% to find the IC_{50} of each sample. Total ammonia in each sample is measured either with the HACH spectrophotometer or an Orion Ammonia electrode. $IC_{50's}$ from the QwikSed tests are compared with total ammonia measured in the leachates and pores to detect confounding toxicity influences.

Solutions of pore waters or leachates are prepared with an enriched seawater medium (ESM) and dinoflagellates at a concentration of approximately 200 cells/mL. Three mL aliquots from each test concentration (ESM volume, pore water or leachate volume, dinoflagellate cell stock volume) are dispensed into five replicates for each test concentration and controls. Cells are cultured directly in disposable optical-grade spectrophotometric plastic cuvettes at 19°C under a light intensity of 4000 lux. Bioluminescence measurements are conducted in consecutive 24-hour increments following test setup.

QwikLite is capable of measuring a response within 24 hours of test setup and can be conducted for a standard 4-day acute test or 7-day chronic test. Similarly, QwikSed can be used to evaluate sediment toxicity and if the contaminated sediment is found to be toxic and needs cleanup by remediation processes, QwikSed can be used to assess the toxicity reduction process.

Key Design Criteria:

- Horizontally-mounted, 2-inch RCA 8575 PMT with S-20 response
- QwikSed controller for processing and counting
- Direct-drive stirring system with adjustable speed motor that uses a stainless steel shaft terminating in a plastic propeller
- Fully adjustable timer and automated data acquisition cycle

2.2 Strengths, Advantages, and Weaknesses.
There are far more sensitive and accurate methods for measuring contaminants in the laboratory as compared to these screening tools. These laboratory methods, however, are slow, laborious and expensive. The field screening tools allow for the rapid mapping and ranking of contaminated sites. With the guidance of this low-cost tool, a high density of semi-quantitative data can be generated on site in near real time. By pinpointing hotspots and ranking relative contamination levels, these data can guide further sampling, and an intelligent selection of meaningful, rather than random, samples for subsequent, more quantitative laboratory analysis. Additional limitations and artifacts affecting the performance of the screening technologies are described in section 5.

By combining FPXRF, UVF, and QwikSed it will be possible to screen for multiple contaminants and their possible biological effects in a more cost-effective manner. Integrating a number of screening techniques creates a further advantage in that using three techniques at a site will facilitate a more efficient and comprehensive mapping of the extent of contamination. Combining screening data with a selected number of laboratory analyses to fully characterize the nature of possible contamination will result in the most efficient analysis plan to characterize the nature and extent of contamination at a site. Specific advantages of the individual technologies are further described in the following paragraphs.

2.2.1 FPXRF.
This screening tool enables the detection of multiple metals in minutes without destroying the sample. Moreover, the FPXRF measures a wide range of elements (Sulfur through Uranium) simultaneously at concentrations between a ppm and percent levels. FPXRF is chosen for its

extraordinary sensitivity, high specificity, simplicity, and low cost. It is a widely accepted, powerful technique that is used for a variety of environmental, industrial, and biotechnology applications. FPXRF is a relatively simple analytical technique that involves minimum sample handling. FPXRF's sensitivity and specificity reduce or eliminate the sample preparation procedures often required to concentrate analytes or remove interferences from samples before analysis. This reduction in or elimination of sample preparation time not only simplifies, but also expedites the analysis.

The principal limitations of this technique are that it is 1) matrix sensitive, 2) semi-quantitative, and 3) is elemental, rather than species-of-molecule specific.

There are several types of interferences that can affect FPXRF performance. These include physical matrix effects, chemical matrix effects, and moisture content. Although it is difficult to eliminate all of these interferences, their effects can be minimized (see section 5 for discussion).

Because of the principal of XRF operation (*i.e*. surface measurement), this technique is not able to provide absolute quantitation of metals in soils or sediments. Other standard methods such as Flame Atomic Absorption Spectrometry or Inductively Coupled Plasma Atomic Absorption Spectrometry, which involve digestion of the sample, are more appropriate for absolute quantitation.

Finally, FPXRF provides a bulk measurement of a metal concentration. It is not capable of differentiating among metal species (i.e., cannot differentiate between Cr^{+3} v. Cr^{+6}), or molecules (*e.g*. [Cl], not [PCB]).

2.2.2 UVF.
This UVF field screening method is used to rapidly determine the location and relative extent of PAH contamination in sediment. As with FPXRF, the method yields qualitative and semi-quantitative results, making it appropriate for preliminary assessments of contaminant distribution as in environmental field screening applications. The high sensitivity and ease of operation of a field fluorometer make fluorescence the method of choice for field screening. UVF utilizes solvent extractions of the bulk sediment to improve PAH detection levels even further, down in the low ppm range. Method sensitivity can vary depending on a number of factors including: sediment matrix, extraction solvent, excitation and emission wavelengths, and specific PAHs present (see Section 5 for additional discussion of limitations).

2.2.3 QwikSed.
Protection of aquatic species requires prevention of unacceptable effects on populations in natural habitats. Toxicity tests are conducted to provide data to predict what changes in viable numbers of individual species might result from similar exposure in the natural habitat. Information might also be obtained on the effects of the material on the health of other species. Bioluminescent dinoflagellates represent an important eucaryotic group, which are widely distributed in the oceanic environment.

QwikSed bioassays can help piece together all of the elements that determine whether or not a targeted area is in need of remediation or control. The use of bioluminescent dinoflagellates, as part of a broader-based biological and chemical testing program, can help identify a potential

problem. By analyzing biological effects using QwikSed, unsuspected contaminants may be indicated.

Substantial savings in operational costs can be achieved by use of this system when compared to other standard bioassays. The QwikSed system can save money in conducting these toxicity tests when compared to conducting the more traditional tests using shrimp, fish, and amphipods. QwikSed requires less time to set up the bioassay and less time to conduct the test.

The QwikSed Bioassay System has been shown to have equal sensitivity to other standard bioassays and can also be used as a mapping tool for determining the extent of marine contamination of sediment pore waters in a fairly short period of time. Without this system, more costly and time consuming methods for toxicity determination of effluents and sediments will be necessary to determine compliance-related issues. Standard bioassays are time consuming to implement (4-8 days of labor per test for an acute 4-day test) and expensive when compared to the proposed QwikSed system (6-7 hours of labor) for conducting an equivalent test.

Ammonia appears to be a confounding influence in QwikSed as well as other sediment bioassays by causing more toxicity than would be displayed in less sensitive organisms (see Section 5 for additional discussion of limitations).

2.3 Factors Influencing Cost and Performance.

The main factors affecting cost and performance of the screening technologies are the number of samples that will be screened and will be sent on for laboratory confirmation. As the number of samples to be screened increases, the per-sample-cost will decrease. The number of samples that go to the confirmatory laboratory can also affect cost and performance. Between 10-50% of the screened samples are usually sent on for costly laboratory confirmation. Site-specific calibration relationships between screening and laboratory data will carry more confidence as the number of laboratory analyses increases, but this will come at a higher cost. All of these factors affect cost and performance, and professional judgment must be exercised to optimize the screening operation (see Chapter 9 in Gilbert, 1987).

3. Site/Facility Description

3.1 Background.

SSC SAN DIEGO searched for suitable demonstration sites with sufficient contamination levels and ranges to demonstrate screening tool capabilities. It was determined that Naval Air Station (NAS) Alameda and the Pearl Harbor Naval Center contain several potential sites with metal and hydrocarbon contamination suitable for demonstrating the XRF, UVF, and QwikSed technologies. These sites were selected based on the following criteria:

- Demonstrations done at the same time as ongoing regulatory projects can offset some of the demonstration/validation costs of the screening project, including ship and sampling operations, laboratory analyses, etc.
- The updated results from screening methods can be used by the regulatory projects, and results will receive wide circulation among regulators and the public.
- The sediment contaminant levels identified during previous investigations ranged from below analytical laboratory detection limits up to greater than significantly high levels (above those causing adverse biological effects (Long *et al*., 1995)). The analytical results from the sites suggest that adequate levels of metals and PAH's exist to demonstrate the XRF, UVF and QwikSed technologies.

3.2 Site/Facility Characteristics.

3.2.1 Naval Air Station Alameda.

Naval Air Station (NAS) Alameda was chosen as the preliminary test site for demonstrating XRF, UVF, and QwikSed technologies. NAS Alameda is located on Alameda Island, at the western end of the City of Alameda in Alameda County, California. Alameda Island lies along the eastern side of San Francisco Bay adjacent to the City of Oakland (Figure 4). The rectangular-shaped base is about 2 miles long and 1 mile wide and occupies 2,634 acres. NAS Alameda includes 1,526 acres of land and 1,108 submarine acres. The majority of the base is land that was created by fill. The perimeter of NAS Alameda bordering open water is retained by a rock seawall. Alameda Island is at the base of a gently westward-sloping plain that extends from the Oakland-Berkeley hills on the east to the shore of the San Francisco Bay. Originally a peninsula, Alameda Island was detached from the mainland in 1876 when a channel was cut linking San Leandro Bay with the San Francisco Bay. The northern portion of Alameda Island was formerly tidelands, marshlands, and sloughs adjacent to the historical San Antonio Channel, now known as the Oakland Inner Harbor. Much of the land that is now NAS Alameda was originally under water. Land use in the vicinity of NAS Alameda is primarily residential and industrial. The base is bordered to the north by the Oakland Inner Harbor, north of which is the main site of the Naval Supply Center Oakland. San Francisco Bay is located west and south of NAS Alameda. To the east is a mixture of industrial, residential, and public land uses including shipyards, naval supply centers, single-family homes, apartments, restaurants, retail stores, schools, and a state beach.

The San Francisco Bay Area experiences a maritime climate with mild summer and winter temperatures. Prevailing winds of the San Francisco Bay Area are from the southwest. Because of the varied topography of this area, climatic conditions vary considerably throughout the region. Heavy fogs occur on an average of 21 days per year. Rainfall occurs primarily during the months of October through April. NAS Alameda averages about 18 inches of rainfall per year (Air Traffic Control, NAS Alameda, 1992). There are no naturally occurring surface streams or ponds on NAS Alameda. Precipitation either returns to the atmosphere by evapotranspiration, runs off into the storm drain system that discharges to San Francisco Bay, collects in artificially created depressions, or infiltrates to the groundwater.

Most of the land at NAS Alameda was created by hydraulically filling existing tidelands, marshlands, and sloughs. The majority of fill used in creating the land at NAS Alameda was dredged material from many areas, including the Oakland Inner Harbor (E&E, 1983). An aerial photograph from 1938 shows a hydraulic dredge, working on the south side of what is now NAS Alameda, placing dredge material in what is now the middle of NAS Alameda (Pacific Aerial Surveys, 1938). Sediments at the surface are coarse and well drained, although there are now no surface channels or drainage sloughs. The original tidal area consisted of deep deposits of Holocene Bay sediments interspersed with numerous drainage channels and sloughs.

NAS Alameda is underlain by about 500 feet of unconsolidated sediments, which themselves overlie consolidated Franciscan bedrock (Radbruch, 1957). The following unconsolidated units are present beneath NAS Alameda and are listed beginning with the deepest (oldest) to most shallow (youngest):

Lower Pleistocene deposits of the Alameda Formation, immediately overlying Franciscan bedrock, consist of undifferentiated terrestrial (channels of sand and gravel, with silt and clay interbeds) and estuarine (relatively finer-grained material containing sparse microscopic marine fossils deposited in bays and marshes) deposits. The Alameda formation ranges from approximately 200 to 400 feet thick beneath NAS Alameda (Rogers and Figuers, 1991).

Upper Pleistocene estuarine deposits, which overlie the Alameda Formation, consist of dark, greenish-gray silty clay. The deposits are considered an aquitard in the NAS Alameda area and are present at a depth of about 90 feet under the westernmost portions of NAS Alameda (PRC, 1991). Upper Pleistocene/Holocene deposits of the Merritt sand differentially overlie the Upper Pleistocene deposits. The unit consists of Eolian (fine-grained sand to silty sand deposited by wind, with bivalve shells and broken shell debris or "hash") deposits which are 0 to 60 feet thick beneath NAS Alameda (PRC, 1996b). Holocene Bay sediments, the youngest naturally occurring unit, consist of fine-grained material (clay to silty clay with silty and clayey sand interbeds, some bivalve shells and plant remains) deposited in an estuarine environment. The Bay sediments are 10 to 110 feet thick beneath NAS Alameda (PRC, 1996b). Artificial fill overlies the Holocene Bay sediments and consists of dredge spoils from surrounding San Francisco Bay, the Seaplane Lagoon, and the Oakland Inner Harbor. The fill composition is generally silty sand to sand with minor inclusions of clay or gravels or both. The thickness of this unit ranges from zero to 30 feet over most of NAS Alameda.

The demonstration project will concentrate efforts in Seaplane Lagoon and deep-water piers on the south side of the lagoon. The lagoon has an area of 110 acres and is located at the

southeastern corner of NAS Alameda. Sea walls surround most of the lagoon, inhibiting the natural flushing processes of bay tides. A breakwater extending from Pier No. 1 forms the southern wall of the lagoon. The entrance to the lagoon is through an 800-ft-long opening in the breakwater. The depth of the lagoon varies from small beach surfaces to a depth of 15 feet. Outside the Seaplane Lagoon are berths for deep draft ships (Piers 1, 2, and 3). These berths are protected by an outer breakwater and have periodic maintenance dredging. No regular dredging program has ever existed at the Seaplane Lagoon, and sediment accumulation is evident in many areas of the lagoon.

Industrial wastewater generated at NAS Alameda before 1974 was discharged directly to the storm drains. The storm drains, in turn, discharged to the Seaplane Lagoon and other offshore areas. The wastewater discharged in the lagoon from 1940 through 1975 was reported to contain heavy metals, solvents, paints, detergents, acids, caustics, mercury, and oil and grease (E&E, 1983). Ship wastewater, which may have contained solvents, chromium, waste oil, and fuel, was also released into the lagoon (E&E, 1983). Between 1972 and 1975, the industrial waste collection system was rerouted to discharge to the East Bay Municipal Utilities District (EBMUD) wastewater system. The Navy now conducts a storm water pollution prevention program to ensure that only rainwater is discharged through the storm drain system. A removal action to remove sediments from the drainage areas of the storm drains was performed in 1995 and the storm drain lines were steam cleaned in November 1996. Other chemicals may have entered the lagoon due to tidal action sweeping ship wastewater—possibly containing solvents, chromium, waste oil, and fuel—from the berthing area into the lagoon. Continuing sources of chemicals may include sediment contamination caused by current berthing practices or historical activities at Piers 1, 2, and 3.

3.2.2 Pearl Harbor Naval Center.

Pearl Harbor is a large complex natural estuary and a major feature located on the south coast of Oahu in the Hawaiian Islands (Figure 5). A majority of Pearl Harbor lies within the Pearl Harbor Naval Center (PHNC). It is located in the southern portion of the Ewa plain, approximately 5.8 miles northwest of downtown Honolulu. Pearl Harbor contains 2,024 hectares (8 square miles [sq mi]; 5,000 acres [ac]) of surface water area and 58 kilometers (36 miles) of linear shoreline. Through the influence of drainage, the Pearl Harbor estuary is the receptacle for runoff from approximately 28,502 hectares (110 sq mi; 70,400 ac) of upland habitat comprising the watershed for much of the southern portion of the island of Oahu.

Grovhoug (1992) provides the following brief history of the PHNC. The PHNC has existed for nearly a hundred years and has undergone extensive changes since the mid-1800s when "Pu'uloa" (as Pearl Harbor was known by the ancient Hawaiians) was a large natural inland lagoon. Numerous walled fishponds located inside the harbor were used to cultivate various species of fish until the 1890s.

As one of the finest natural harbors in the Pacific Basin, Pearl Harbor was readily identified as a strategically important military asset. The U.S. Navy acquired rights to the harbor in an agreement with King David Kalakaua in 1873 (U.S. Department of the Interior, 1969). After 1898, when Hawaii became a territory of the United States, plans were developed to dredge the harbor entrance channel and construct docking facilities inside the harbor. In 1901, the U.S. Navy acquired 800 acres (ac) of land to establish a Naval Station on Pearl Harbor (U.S. Navy,

1983). The first major dredging of the entrance channel began in 1908, followed by construction of the first drydock in Hawaii at the Pearl Harbor Navy Yard (Nystedt, 1977). After problems were encountered with underground water pressure, Dry Dock #1 was finally completed in 1919 (U.S. Navy, 1983).

During World War I, a dozen warships were repaired and overhauled at the Navy Yard. From 1917 to 1918, a temporary submarine base was relocated from Magazine Island (Kuahua Island) to Quarry Point on the eastern shoreline of Southeast Loch. A naval ammunition depot was commissioned in 1919 at Magazine Island. Around 1920, many walled fishponds still remained intact.

During the 1920s and 1930s, shore facility developments continued and additional land was acquired by the Navy. Ford Island (formerly known as Moku'ume'ume, "island of the little goats") became a naval air station in the early 1920s. Work began on concrete moorings along the south side of Ford Island, which later became known as "Battleship Row." Industrial development was greatly accelerated in the Pearl Harbor area during the late 1930s and early 1940s. A considerable amount of acreage in the Pearl Harbor Naval Complex has been created since 1930 by the deposition of dredge spoil materials (U.S. Navy, 1947).

On 7 December 1941, the Japanese Imperial Navy launched a surprise air attack on the U.S. Fleet in Pearl Harbor from a task force of 32 vessels, including 6 aircraft carriers with 350 warplanes. This attack sank or severely damaged 21 of the 86 U.S. Navy warships in Pearl Harbor (Lenihan 1989, U.S. Navy 1989a). Chemical evidence (*i.e*. elevated concentrations of copper, lead, and zinc) of this period remains detectable in buried Middle Loch sediments that have not been disturbed by dredging activities (Ashwood and Olsen, 1988). They also report that the bombing attack resulted in about six times more lead input to this estuarine area than the total combined lead input from sewage disposal and naval maintenance operations during the succeeding 45 years.

From 1940 to 1943, large amounts of dredged material were placed on Waipio Peninsula and areas adjacent to the Submarine Base (U.S. Navy, 1983). These landfill operations formed the present shoreline configuration of the inner harbor. From 1942 to 1944, the number of facilities and personnel at the PHNC increases greatly to support the war in the Pacific. Storage facilities for ordinance and material filled nearly all of the available land regions near Pearl Harbor. By mid-1943, civilian employment at the Navy Yard rose to 24,000 personnel (U.S. Navy, 1983).

After World War II and throughout the late 1940s, the number of service personnel and active facilities at Pearl Harbor decreased markedly. During the Korean War and the Vietnam conflict, operations and support personnel at the PHNC increased in response to the nation's defense requirements, but never to the same extent as during World War II. Today, Pearl Harbor is a major fleet Homeport for nearly 40 warships; service force; vessels and submarines; and associated support, training, and repair facilities. The region is also listed as a National Historic Landmark.

During the last century, many human activities have been concentrated along the shoreline and within the upland drainage basins that empty into the harbor. These activities include the industrial and operational activities of the U.S. Navy; private industrial operations; municipal,

commercial, and urban activities; and agriculture. These activities potentially release numerous types of chemical contaminants into the air, water, and soil along the shoreline and within the drainage basins that empty into Pearl Harbor. The approximately 2,024 hectares (5,000 ac) of soft (*e.g.* mud and sand) sediments comprising the bottom in Pearl Harbor are the ultimate sink or repository for these chemicals and the natural habitat for thousands of estuarine and marine species.

The present day PHNC is an outgrowth of more than 100 years of peacetime and wartime development that has resulted in (1) dredging to construct a channel and berthing area of sufficient depth to allow passage of the "largest of ships" (Grovhoug, 1992) and (2) construction of extensive shoreside facilities (*e.g.* ship mooring and repair facilities, fuel storage, handling, transfer, and recycling facilities as well as operations, maintenance, and support facilities) to meet changing needs of the U.S. Fleet. Military vessels using the harbor on a regular basis include U.S. Navy surface ships, submarines and harbor craft; U.S. Army cargo transport vessels; U.S. Coast Guard buoy tenders and patrol vessels; U.S. Coast Guard and patrol vessels; and foreign naval vessels. Harbor navigation channels and mooring areas at piers and wharves supporting these vessels are maintained at water depths necessary for safe navigation through a program of routine maintenance dredging. New facilities are developed as needed and may involve in-water construction and project specific dredging.

Mean annual rainfall in the vicinity of the PHNC is approximately 64.8 centimeters (cm) (25.5 inches). The PHNC is relatively dry when compared with other areas on Oahu, particularly just leeward of the crest of the Koolau Range where mean annual rainfall may exceed 275 inches. Rainfall is seasonal, varying from 10.2 cm (4 inches) per month during the winter (December to February) to 2.54 cm (1 inch) per month during the summer (June to July) (Giambelluca *et al.,* 1986).

The prevalent winds across the PHNC are the northeast trade winds that prevail for approximately 9 months of the year. The mean wind speed is 11.6 miles per hour (mph). During the balance of the year, south to southeast winds and mild offshore breezes prevail. The south winds are usually accompanied by wet tropical air and frequent showers. During the summer months, periods of "no wind" occasionally occur but do not persist for more than a few days.

Temperature varies considerably by season as well as diurnally in the Pearl Harbor region. During the summer months, afternoon high temperatures range between 30.5 and 31.6 degrees Celsius (^{o}C) (87 and 89 degrees Fahrenheit [^{o}F]), and nighttime low temperatures range between 22.2 and 24.4 ^{o}C (72 and 76 ^{o}F). In the winter months afternoon high temperatures range from 24.4 to 25.5 ^{o}C (76 and 78 ^{o}F) with nighttime low temperatures ranging from 12.8 to 18.3 ^{o}C (55 to 65 ^{o}F).

Grovhoug (1992) reviewed past environmental information on Pearl Harbor compiled from numerous studies conducted over several years. The majority of these studies are project specific and address environmental concerns at specific locations in the harbor. Some studies were spatially comprehensive providing data for large areas of the harbor (*e.g.* the IAS study of historic contamination, the Evans *et al.* 1974 assessment of biological and physical conditions in the harbor, and the Youngberg 1973a study of metals in the harbor). In general, these studies provide useful background information but are limited for purposes of a harbor-wide assessment

because of their age (*e.g.* some are 20 years old), or they are fragmented over temporal and/or spatial scales.

Our cursory overview of the available data determined a few specific areas of interest to demonstrate the screening technologies. These included the Middle Loch and Bishop Point areas. From this overview, it appears as if the Middle Loch area is very fine-rich (75 – 90%) although it has low TOC values (1.98 – 3.83%). The Bishop Point area appears to be less fines-rich (41 – 56%), yet the TOC values are higher (4 – 6%) than the Middle Loch area. The Bishop Point area is a small pier area (~3 acres) and is rumored to be very heterogeneous, with coral hard bottom to soft mud conditions (pers. comm. Jeff Grovhoug). The pier area is in current use with a number of ships present at any particular time. The Middle Loch area, on the other hand, is very large and more homogeneously fine grained mud. This area is regularly dredged to maintain a draft of 20 ft and used to store a "mothballed" fleet of ships.

The contaminants of concern in these two areas differ. The metals levels in the Middle Loch area are elevated, whereas they are very low at Bishop Point. However, the PAH level at Bishop Point are elevated and range from ~20 – 40 ppm tPAH. PAHs do not appear to be elevated in the Middle Loch area. Of the other contaminants of concern (pesticides, PCBs, TBTs), it appears as if these areas are not very contaminated.

4. Demonstration Approach

4.1 Performance Objectives.
The performance objectives are a critical component of the demonstration plan. They provide the basis for evaluating the performance and costs of the technology. As stated in Section 1.3, the primary objectives of this demonstration are to evaluate the three field screening technologies in the following areas: (1) their performance compared to conventional analytical methods; (2) data quality; (3) the logistical and economic resources necessary to operate the technologies; and (4) the range of usefulness in which the technologies can be operated and integrated into a screening procedure that allows more efficient assessment of sediment sites. Secondary objectives for this demonstration are to evaluate the technologies for their reliability, ruggedness, and ease of operation.

Performance relative to conventional laboratory methods will be evaluated in a number of ways. A plot of screening versus laboratory data will be generated from the site samples. Figure 6 shows such a plot for the UVF pre-demonstration data collected at Alameda. Limits of detection (LOD) for the screening technique are determined from replicate analysis of "clean" samples and marked on the plot by a dashed vertical line at about 250 fluorescence intensity units. A regulatory action level (ERL (Long *et al*., 1995) for Total PAHs) is marked on the plot by a horizontal dashed line at about 4000 ppb. These dashed lines divide the plot up into quadrants representing non-detects, detections above a regulatory action level, false negatives, and false positives. The main performance criteria to be used for comparison of the individual screening techniques to laboratory data are percent false negatives (%FN), correlation coefficients (r^2), and relative standard deviations (RSD). Following the example of USEPA procedures (PRC, 1995), screening data will be classed into three levels depending on these criteria. Level 3 definitive data (r^2=0.85 to 1.00, RSD <10%, %FN <2%) can be considered to substitute for laboratory data. Level 2 Semi-quantitative screening data (r^2=0.6 to 0.85, RSD =10-20%, %FN <5%) require a limited number of confirmatory samples (usually around 10%) for calibration to be considered quantitative. Level 1 qualitative screening data (r^2<0.6, RSD > 20%, %FN <10%) detect the presence or absence of some parameter, but may not quantify concentration levels. Although most screening data is classed as Level 1, the goal of this project will be to demonstrate the data will meet or exceed Level 2 requirements.

It should be pointed out that with QwikSed, which is a screening bioassay, we will not follow the exact USEPA Level 1, 2, and 3 criteria defined for chemical screening techniques. There is a strong positive correlation between QwikSed and the sea urchin development tests ($r = 0.812$; $p < 0.001$) from pre-demonstration data. Both toxicity tests are sensitive to ammonia, a common component in sediment pore waters or leachates. Even with this limitation and confounding influence, there is still a good correlation. Because of the reasonably good correlation between QwikSed results and the laboratory bioassay tests, we will use the laboratory bioassay test as "the standard" to which to compare QwikSed test results against. Statistically significant correlations will be the acceptable criteria to accept performance, but for convenience and planning, correlation coefficients (r) values of 0.6 and greater with $p < 0.05$ should be considered as an acceptable and satisfactory criteria for which to accept QwikSed results.

Part of the data quality objective will be addressed by examining precision and accuracy of the data. As mentioned above, RSDs will be used to judge precision. With 10% of the samples, replicate measurements for each technology will be made to calculate RSDs. Included with these 10% of the samples undergoing replicate analysis will be at least one performance evaluation (PE) sample (typically a NIST level standard reference material but may include a well-characterized internal laboratory standard). As mentioned above, correlation coefficients will be used to judge accuracy. The costs associated with the demo will be tracked and a table generated to document the cost of operating the technologies in the field. Additional qualitative information on the reliability, ruggedness, and ease of operation of the techniques will also be tabled for comparison.

The final primary objective, to judge the usefulness of an integrated screening procedure for sediment assessments, also requires subjective or qualitative evaluation. In one area of study where there will be sufficient demo and pre-demonstration data from both screening and laboratory analyses, a simple comparison will be made between screening and laboratory data. If only 10-20% of the laboratory data are used to calibrate the screening data, it will be determined if the areas of concern picked from the calibrated screening data change when the remaining laboratory data are considered. This will be done for each individual screening technique by comparisons to laboratory data as stated above (using r^2, RSD, and %FN), and with the screening data taken as an integrated process by comparisons of "areas of concern" defined by screening and laboratory data. An "area of concern" is usually defined by regulatory projects through a weight of evidence approach requiring both chemical and biological "hits." If a conservative screening decision rule is used that states any "hit" from an individual screening technique is sufficient to trigger a screening "area of concern," screening should find all "areas of concern". For example, from our total number of sites with both screening and laboratory data, we will show different scenarios for how a limited number of screened samples could have been sent on to the laboratory to define the same "areas of concern" at substantial cost savings. These cost savings could also be used to provide additional sites for screening to reduce the uncertainty of regulatory decisions through greater spatial coverage. Gilbert (1987) provides specific examples of using screening and laboratory data together to optimize for the reductions in costs or data variability depending on the r^2 and cost differential between screening and laboratory data.

During pre-demonstration work, between 20-30 samples were analyzed by screening and laboratory confirmatory techniques. This number appears acceptable to judge the performance of screening techniques compared to confirmatory laboratory results. For this demonstration, the level of data quality to be achieved will be comparable to Screening Data with Definitive Confirmation as described in the Data Quality Objectives Process for Superfund – Interim Final Guidance (USEPA, 1993). According to these guidelines, screening data are those data generated by rapid, less precise methods of analysis with less rigorous sample preparation. Screening data provide analyte identification and quantification, although quantification may be relatively imprecise. Our ultimate goal in producing these demonstration databases is to convince regulators that these screening data reach Level 2 criteria discussed above, and only a percentage (usually around 10%) of the screening data must be confirmed using laboratory analytical methods. For this demonstration, however, all 20-30 samples that will be used for the regulatory project will have screening and laboratory analyses to build our database. According to standard screening procedure, at least three screening samples reported above the action level (if any) and three screening samples reported below the action level (or as non-detects) should be

randomly selected from the appropriate group and confirmed (USEPA, 1993). These guidelines in conjunction with USEPA Method 6200 (USEPA, 1998), which requires that confirmatory samples selected from the lower, middle and upper range of measured concentrations will be followed for this demonstration. These data will be combined with pre-demonstration data from Alameda so a total of over 50 samples should be available to evaluate each screening tool.

4.2 Physical Setup and Operation.
The details of the methodology for the various screening techniques are adapted from standard protocols. These screening techniques have been adapted from USEPA (PRC 1995, Filkins, 1992) or ASTM (Lapota *et al*., 1997) methods. Sediment samples will be obtained by standard grabs or cores. Representative sample splits will be separated for screening and laboratory analyses. Due to the different analysis times required by the various techniques, it is expected that results will be available from FPXRF after several minutes, followed by UVF after half an hour, and finally by QwikSed after 4 to 24 hours.

With the differing analysis times, results will be available from the various screening techniques at different times during the sampling process. We will therefore be depending more heavily on near real-time chemistry results to help guide subsequent sampling locations. The general procedure in mapping out contaminant plumes will be to start at suspected sources (for example, industrial outfall pipes) and work outward to delineate the extent of contamination. If no contamination is detected at the source using one or several screened samples, there will be no need to continue sampling away from the source. If contamination is detected at the source area, sampling continues outward to define the edge of the contaminant plume. Since the biological effects results from QwikSed will not be available until much later (4 to 24 hours), these data will not be available for near real-time guidance during sampling. They will, however, be used together with the chemistry screening data to select which samples continue to the laboratory for full characterization. It is anticipated that the laboratory samples will be selected to span the full range of results observed in all screening techniques. This will allow calibration curves between screening and laboratory techniques to cover the entire range of observed results and therefore allow better predictions from the remaining screening results.

4.3 Sampling Procedures.
Site contractors (PRC and site sub-contractors) will be conducting sampling and analysis for the regulatory project, so they will be handling all site setup and facilities. SSC SAN DIEGO will provide FPXRF, UVF, and QwikSed equipment and operators who will recover a sample split for screening analysis. Remaining sample will continue on to the laboratory for confirmatory analyses. The sampling plan for this demonstration of the FPXRF, UVF and QwikSed technologies specify procedures that will be used to ensure the consistency and integrity of samples. In addition, this plan outlines the sample collection procedures necessary to meet the demonstration purpose and objectives.

4.4 Analytical Procedures.
To assess the comparability of the data acquired by the FPXRF, UVF, and QwikSed screening technologies to data generated by established, conventional analytical methods, the screening data will be compared to confirmatory analysis results. The overall objective of the sampling program is to collect FPXRF, UVF and QwikSed and traditional analytical data in parallel to

demonstrate the FPXRF, UVF, and QwikSed technologies' capability to delineate the extent of sediment contamination.

The selection of analytical laboratory and methods, plus sample collection and analysis procedures, will be documented in a separate sample and analysis plan document prepared by the Navy site contractor, PRC management Inc.

5. Performance Assessment

5.1 Performance Data.

5.1.1 FPXRF Analytical Method.

Elemental analyses for this demonstration were performed using a TN Spectrace 9000 portable XRF spectrometer (TN Spectrace Instruments) (Figure 7). The instrument contains three radioisotope sources, Fe-55, Cd-109 and Am-241 to provide the excitation x-rays. It has an electronically cooled solid-state mercury iodide detector for measurement of the characteristic fluorescent x-rays. The Spectrace 9000 FPXRF is capable of measuring up to 25 elements, ranging from sulfur to uranium, simultaneously. The instrument was calibrated at the factory using pure elements and proprietary fundamental parameters (FP) algorithms, which eliminate the need for empirical calibration with site-specific standards. One such FP algorithm, the "coarse-grain" application has additional correction factors for grain size effects and is used primarily for wet, heterogeneous sediment samples. Contaminant concentrations are computed using the matrix-correcting FP algorithms. Typically, count times of 200-sec (Cd-109), 20 sec (Fe-55) and 20 sec (Am-241) are used for analysis.

The coarse-grain FP application was used for all pre-demonstration and demonstration sample analyses. Approximately 10 grams of wet sediment were used for each sample analysis. Only one operator was required for sample preparation and analysis. Four metals, Fe, Cu, Pb, and Zn, were measured by FPXRF for this demonstration. Copper, lead and zinc are common contaminants at military sites and are found in marine sediment at levels detectable by FPXRF. Other elements of concern (*e.g*. As, Cd, Cr, Hg, Ni) were not measured because they are often present in marine sediment at concentrations below method detection limits. While absolute concentrations of metals in sediments are an important part of assessing a site, there are a number of reasons that this alone does not provide a full picture of what is going on at the site. Both organic and inorganic contaminants can exist in a region at background, ambient or natural levels, either because they have natural sources or because entire regions in urbanized, industrialized and other areas are exposed to ubiquitous levels of anthropogenic input. Metals, in particular, exist at natural levels in source minerals that generate the sediment matrix. In many cases, since contaminants have a tendency to associate with fine-grained sediments, there is a general regional tendency to have a "mixing curve" of contaminated fines, and relatively uncontaminated coarse-grained sediments.

Often, either ambient contaminant levels or background natural levels or a combination of both can be separated from site-specific levels by normalizing to or plotting against sediment characteristics which tend to indicate natural metal-rich particles (e.g., Fe, Al) or fine-grained particles (e.g., Fe, Al, %fines, %OC). Thus, it is important at a given site to examine contaminant distribution relative to regional ambient or background levels. To address these issues, Fe was measured in the demonstration samples.

Several factors can affect FPXRF performance. Performance is primarily affected by operating conditions and matrix interferences. Some sources of interference can be minimized or

controlled, while others cannot. The effect of many of these factors on the demonstration results will be discussed.

Operator related factors are the most easily controlled. Differences in operator technique can have a significant effect on numerical results. For FPXRF technologies, variation in sample preparation and measurement technique by the operator can affect the results. These effects can be controlled through the use of the same personnel to prepare the samples and to operate the instrument throughout the demonstration or by careful training of an alternate analyst. Performance evaluation results also signify operator-dependent errors. These errors were controlled by having one experienced analyst carry out all demonstration and pre-demonstration analyses, as was done during the demonstrations.

Another factor, which affects results, is sample analysis time. Not only does analysis time affect sample throughput, it affects precision and detection limits. Increasing the count time by a factor of four will improve precision by a factor of two and improve detection limits by 50%. However, there is a point of diminishing return. It is impractical in terms of sample throughput as well as non-beneficial in terms of improved precision and detection limits to extend source count times beyond 600 to 800 seconds. This factor can be controlled by careful and consistent adherence to specific SOPs.

Physical matrix effects result from variations in the physical character of the sample. These variations can include such parameters as particle size, uniformity, homogeneity, and surface conditions. There are fundamental differences between the way FPXRF and standard analyses treat and measure a sample that limit the degree of direct comparison of the results. In standard analyses, a sample is either partially or completely digested; the extract is cleaned up and analyzed. FPXRF, on the other hand, is a bulk surface analytical method in which x-rays bombard the surface of a sample, exciting fluorescence in that portion of the sample (~1 mm) that the x-rays penetrate. Thus, it can be more sensitive to sample heterogeneity than is a standard digestion analysis.

Digestion of a sample allows for analysis of all the extractable metal in the sample, whereas FPXRF analyzes the metals near the sample surface. If a highly contaminated or pure metal particle that drives the total concentration in that sample is in the sample but not at the surface, digestion will reflect the total concentration in that sample, whereas FPXRF will reflect the concentration in the exposed cross-section. On the other hand, if a high concentration particle is at the surface and accessible to FPXRF, but is not at high enough levels to dominate the bulk concentration, FPXRF may overestimate the total concentration. Although it is impossible to completely overcome the effects of physical matrix interference, measures such as the analysis of field duplicate samples and thorough sample homogenization prior to subsampling and analysis can be used to minimize the effects. However, it should be pointed out that at very heterogeneous sites, or those at which contaminants of concern are associated with large, randomly distributed particles, heterogeneity is a problem encountered by "standard" extraction-based methods as well. With such samples, large error bars, variability between field duplicates or subsamples, and non-reproducible results are a common problem. If the results of these standard analyses are designated as "truth" against which results from another method is compared, that second method appears inaccurate. However, in heterogeneous samples, no "truth" exists, inasmuch as a given sample only represents itself, not an area.

Moisture content may also affect the accuracy of sediment and soil analyses. Differences in the way FPXRF and extractive analyses are carried out make FPXRF more sensitive to moisture content than are standard methods. In an extractive analysis, analyte concentration can be normalized to dry weight, since samples can be dried and weighed either before or after extraction. Thus, variations in water content do not affect the results. However, if a sample is directly analyzed in the field, there is no dry weight measurement, and moisture content cannot be corrected for. Since the FPXRF measures a bulk concentration of a sample's surface cross section, this measurement is sensitive to water content in that cross section. Particularly in a wet sediment sample, some settling will occur and the water content in the cross section analyzed at the bottom of the cup may not be representative of the total sample water content. If a sample is dried and ground, matrix effects related to sample heterogeneity and moisture content are decreased. Furthermore, with the intervening water removed, the bulk metal concentrations are higher, and thus the FPXRF is more sensitive to the metals in these samples. Stallard *et al*. (1995) found, however, that although drying and grinding samples before FPXRF analysis did help improve the accuracy and sensitivity of the measurement, the offset between wet and dry measurements was not always directly proportional to sample water content, since differential settling in the sample cups as a function of grain size can also cause variability not observed in an extractive analytical method. Sediment samples were not dried and ground during the demonstrations, since this process detracts from the utility of the instrument for rapid on-site screening, which is the purpose of these demonstrations. The ability of the FPXRF to produce a fairly consistent response regardless of multiple variables such as moisture content, heterogeneity, mode and source of contamination and sediment mineralogy has been demonstrated (Kirtay *et al*., 1997).

Finally, chemical matrix effects result from interactions between fluorescent x-rays from different elements. These effects can occur as either spectral interferences (peak overlaps) or as X-ray absorption and enhancement phenomena. Both effects are common in sediments and soils contaminated with heavy metals. Due to the complexity of these interferences, they will not be discussed here. However, a thorough discussion of these interferences can be found in EPA Method 6200 (EPA, 1998). Several of these effects can be corrected through the use of Fundamental Parameter (FP) coefficients. Calibration by Fundamental Parameters was used throughout the demonstration. Where observed, these effects are reported in the data analysis.

Understanding the factors that can affect FPXRF performance will allow for an evaluation and utilization of those options necessary to minimize the effects and meet the required objectives.

In support of various assessment projects, the FPXRF was deployed on site at NAS Alameda on several occasions. Confirmatory analyses were carried out for two of these deployments. The results from these two deployments are presented here. The first set of results is from a pre-demonstration deployment (07/97) and the second set of results is from the demonstration deployment (10/98). The samples for the demonstration were collected in conjunction with an on-going regulatory project at the site. During the pre-demonstration, 31 sediment samples were collected and analyzed on site by FPXRF. Fifteen of these samples were sent out for confirmatory analyses. During the demonstration, 29 sediment samples were collected. During the first two days of deployment, 11 samples were collected from Seaplane Lagoon and were analyzed on site. The remaining 18 samples were collected the following week, and were subsequently analyzed in the laboratory. The 11 samples, which were analyzed in the field, were

also re-analyzed in the laboratory. All 29 samples were sent out for confirmatory analyses as required by the regulatory project. Because the purpose of this report is to validate on site screening tools, for the bulk of the data discussion, only the results from the analyses performed on site (pre-demonstration and demonstration) are presented. However, on site and in-laboratory FPXRF results will be compared and briefly discussed. Furthermore, in-laboratory FPXRF results will be compared to certified results.

For each sample collected, a sample split was prepared, packed on ice and sent to a laboratory for confirmatory analysis by standard methods (Fe, Cu and Pb were analyzed by Flame Atomic Absorption Spectrometry (FAAS); Pb was analyzed by Inductively Coupled Plasma Mass Spectrometry (ICP-MS)). A portion of the samples (50%) collected during the pre-demonstration deployment was sent to the Florida Institute of Technology for analysis. The samples collected during the demonstration were sent to a CLP laboratory for analysis as required by the regulatory project. For each deployment, a standard report was generated. These reports include the following information: sample ID, sample analysis date, FPXRF results, confirmatory results and QA/QC data. The reports generated for this project have been archived and are available upon request. For the most part, standard operating procedures for the FPXRF (as described in section 5.1.1.1) were followed during each deployment. However, during the demonstration, the count time for the Cd-109 source was increased to 300 seconds to adjust for source decay. No problems were encountered during either the pre-demonstration deployment or demonstration deployment. Table 4 and Table 5 provide condensed versions of the standard data reports. A detailed comparison between the XRF results and certified results is presented in Sections 5.2 and 5.3 (Data Assessment and Technology Comparison).

As described above, both physical and chemical matrix interference effects were expected with samples from NAS Alameda. The anticipated physical matrix effects included moisture and sample heterogeneity effects. The samples collected from NAS Alameda ranged in percent moisture from 30% to 70%. The sediment ranged from coarse, sandy sediment to fine-grained sediment. Sediment samples from certain locations (e.g., samples from the corners of Seaplane Lagoon, samples collected near piers) contained chunks of foreign material including wood from pier pilings, paint chips, and other unidentified particles. The coarse-grained material in most other samples appeared to be primarily shell hash and mineral material. Chemical matrix effects were encountered, which were caused by elevated concentrations of Fe (~ 5%), resulting in an absorption effect of the Cu x-rays, thereby reducing the intensity of the Cu measured by the detector. Although not all types of interferences can be corrected for, the use of Fundamental Parameter (FP) coefficients can correct for both physical and chemical matrix effects to some degree. Standardless FP calibration was used throughout the demonstration

For pre-demonstration purposes, prior to deployment at the Pearl Harbor Naval Complex, a subset of sediment samples (n=42) collected for a site assessment being carried out in 5 regions of the Naval Complex was obtained for analysis. A split from each of the archived samples was shipped to SSC San Diego by Ogden Environmental for analysis by FPXRF. The sediment samples were prepared according to standard procedure and analyzed in the laboratory.

The FPXRF results from the pre-demonstration samples were compared to certified results. Acceptable correlations between the FPXRF and confirmatory analyses were observed for Cu, Zn, Pb and Fe, although the average slope for these comparisons was 0.3, while previous work

generally resulted in average slopes of about 0.5. The mineralogical composition of the sediment in the Pearl Harbor area (adjacent to a volcanic island) differs from typical continental United States west coast sediments. Sediment samples from the loch regions contain very high levels of iron (5– 15%). The elevated levels of iron were considered to be a cause of spectral interference and thought to affect the ability of the instrument to detect copper and possibly zinc. However, modifications to the methodology were not made, since one goal was to allow for an evaluation of instrument response to different sediment types.

For the deployment, the FPXRF was packed and shipped to Hawaii. The instrument was set up in a makeshift laboratory (Figure 8). After setup, standard instrument check procedures were performed. During this period, it was noted that the instrument was not responding according to the SOP specifications. It was determined that the bias battery that powers the probe was dead. A new bias battery was installed. In accordance with SOPs, a spectral resolution test was performed. Results from the test indicated that the Fe intensity was below the accepted threshold for analysis (0.89 versus 0.95). Repeated tests produced the same result. These results indicated that instrument spectral resolution was diminished. Diminished resolution will affect instrument sensitivity, as well as the ability to accurately identify elements. Nevertheless, all of the samples collected during the deployment were analyzed on site for demonstration purposes. Upon return, the instrument was shipped to the manufacturer for diagnostic testing. No problems with the detector system could be found. The Cd-109 source was replaced and the software was upgraded for Y2K compliancy. Because of concerns over the reliability of the data collected during the deployment, the sediment samples were reanalyzed in the laboratory subsequent to the diagnostic testing and source replacement. These results are used here.

Again, for each set of analyses (pre-deployment and deployment), a standard report was generated. The reports include the following information: sample ID, sample analysis date, FPXRF results, confirmatory results and QA/QC data. The reports generated for this project have been archived and are available upon request. Table 6 through Table 9 provide a condensed version of the standard reports. The same notations used above apply to these tables as well.

In conjunction with this demonstration, a bench-top Energy Dispersive X-ray Fluorescence Spectrometer (EDXRF) was also tested. The samples collected during the pre-demonstration and the demonstration were analyzed in the laboratory using a QuanX EDXRF Spectrometer (Spectrace Instruments, Sunnyvale, CA). This instrument contains a Rh-anode x-ray tube for primary generation of x-rays (4-50 kV) and a thermoelectrically cooled, solid-state Silicon (Lithium) detector. The Si(Li) detector provides spectral resolution which exceeds other solid-state detectors or gas-filled proportional detectors. For this reason, the EDXRF was tested in order to determine if elements could be measured in a screening mode (wet) with increased sensitivity. Although the size of the instrument precludes its use as a field instrument, rapid turnaround of results is still possible. Samples can be shipped overnight to the laboratory and analyzed the following day (see Addendum).

5.1.2 UVF Analytical Method.

Analyses for polycyclic aromatic hydrocarbons (PAHs) for this demonstration were made using a Turner Designs Field Fluorometer Model TD-10 (Turner Instruments) (Figure 2). The instrument is fully operational in the field with either battery or standard 110V power. It has several standard optical packages (source lamps and optical filters, see Chapter 2 for additional instrument description) designed specifically for measuring PAHs. For this demonstration 4-gram sediment samples were placed in 50 ml disposable centrifuge tubes and mixed with varying amounts of a drying agent (5-10 grams Sodium Sulfate powder). After mixing for several minutes to facilitate drying, 20 ml of hexane solvent was added to each tube. Individual tubes were placed in a vortex mixer for 30 seconds to ensure solvent was mixed into solid sample, and then groups of tubes (usually batches of 10 –20 tubes) were placed on a shaker table for 20 minutes for an extraction step. Tubes were then placed on a low speed centrifuge to separate the solvent from the solid sample matrix, and the hexane solvent (with dissolved PAHs) was pipeted into glass test tubes. With the low wavelength optical package (excitation at 254 nm), quartz test tubes are required, whereas with the high wavelength optical package (excitation at 350 nm) standard disposable glass test tubes were used.

As with FPXRF, many factors can affect performance. Like in the XRF section above, these will be discussed in an order from the more easily controlled operator factors to the more difficult to handle matrix variables. Differences in operator technique can have a significant effect on results. For UVF, variation in sample preparation is more important than fluorescence variability because of the high repeatability of the fluorescence measurement on homogeneous liquid extracts. Since the PAHs are extracted with a hexane solvent, variation in the extraction step (time and degree of mixing) can lead to variable measurement results. All samples should therefore be handled in a consistent manner to ensure uniform extractions. This variability can easily be controlled through the use of the same personnel to prepare the samples throughout the demonstration or by careful training of an alternate analyst. For this demonstration these errors were controlled by having one experienced analyst carry out all demonstration and pre-demonstration analyses in a uniform manner.

It should be noted that lab analyses were also conducted with an extraction step, but rather than hexane they used dichloromethane (DCM) as the solvent for extraction. DCM is a stronger extraction solvent, but doesn't have the required optical properties for the fluorescence measurements. Hexane is optically transparent at the required wavelengths for the fluorescence measurements, but results in a somewhat lower concentration of PAHs in the solvent extracts (all other factors (time, mixing degree, etc) being equal). Early experiments carried with varying extraction procedures (varying mixing times with shaker table, vortex mixing, and sonication) were employed to select the utilized extraction procedure. Experiments on a limited number of samples indicate these various solvent extractions (hexane, acetonitrile, DCM, etc.) result in proportional concentrations of PAHs, with DCM more concentrated than hexane. It should also be noted that the lab analyses, treated here as the absolute "true" PAH value, are really just the operational defined "true" values. Surrogate PAHs spiked into laboratory samples before DCM extraction often only show a percentage of the spiked PAH in the measured GC/MS value. Although not done in this demonstration, similar spiking with surrogates could be done to check for consistent proportions of PAHs in hexane extracts for fluorescence. It is likely that variations in sediment matrix (grain size, TOC, particle composition, etc.) and extraction solvents do result in variations in extraction efficiency, but this effect on fluorescence is probably lost in the

"noise" of other variables. For this demonstration it was assumed that PAHs in hexane extractions were proportional to bulk sediment PAH concentrations, regardless of sediment matrix. The fluorescence values are then "calibrated" to a set of lab analyses that use a specific extraction procedure, assuming the PAH concentrations in the hexane extracts are proportional to the PAH concentrations in the extracts used in the lab analyses.

Another matrix factor that plays a large role in variation in fluorescence values is the mix of individual PAHs that are present in the sample. In addition to the above-mentioned differences in extraction efficiencies, different PAH mixtures result in different fluorescence characteristics. The total PAH mixture is composed of many more individual PAHs than the standard 16 parent compounds normally reported in the standard EPA method 8270 analyses. Some alkylated forms of these 16 parent compounds actually show more fluorescence than their parent compounds. This leads to site specific fluorescence responses depending on the mix of PAHs that are present at the site. For this demonstration we can account for this variability by using different calibration curves depending on the site. Owen *et al*. (1995) looked at the effect of various calibration methods for PAH screening by fluorescence, including the benefit of running site specific splits versus standard reference materials (SRMs). This factor leads to the recommendation that a certain percentage (10-50%) of screened samples should always have splits sent for confirmatory lab analyses. This will allow for site specific calibration curves to be generated to convert raw fluorescence values into corresponding PAH values.

5.1.3 QwikSed Analytical Method.

A series of field exercises were conducted which compared the performance of QwikSed against other more traditional sediment toxicity tests. Sediment samples, following collection, were either tested on-site or shipped to another laboratory over night to conduct another toxicity. These other tests included the sea urchin development test (72-hr test) or the amphipod survival (10 day) toxicity test. The first demonstration was conducted at NAS Alameda with QwikSed and the sea urchin development test. Some pre-demonstration amphipod survival data at the same sampling locations was also included in the demonstration data. The second site QwikSed was demonstrated was at Pearl Harbor, Hawaii. The areas surveyed were Middle Loch and Bishop Point. Sediment was shipped overnight to MEC Laboratories in Carlsbad, CA where the sea urchin development test was conducted. Toxicity thresholds for all tests were established from earlier studies. Samples with less than 84% of control cell values were considered toxic. Samples that had less than 50% amphipod survival were considered toxic. Samples which resulted in less than 80% normal larval development were considered toxic.

SEDIMENT ELUTRIATES

Sediment elutriates were prepared from all samples following refrigeration of the samples at 4°C. The elutriates were prepared by mixing sediment with filtered seawater in a 1:4 ratio, mixing for 30 minutes, settling for 1 hour, and pouring off the seawater (EPA-503/8-91/001, Feb 1991). The elutriate is then filtered through an 8 μm filter membrane. Five elutriate concentrations were prepared to identify the level of toxicity: 6.25%, 12.5%, 25%, 50%, and 100%. Control cells were also tested alongside each batch of prepared sediment elutriate cells (ASTM E 1924-97, Standard Guide for Conducting Toxicity Tests with Bioluminescent Dinoflagellates; published June 1998). Enriched seawater medium (ESM) was used as makeup water (ASTM E 1218-90).

BIOLUMINESCENT SPECIES TEST ORGANISMS

In these field demonstrations, 2 species of bioluminescent dinoflagellates were used as the QwikSed test organism: either *Gonyaulax polyedra* or *Ceratocorys horrida*. Stock flasks were inoculated with either of these dinoflagellates to obtain a concentration of 200 or 100 cells ml^{-1} of ESM and elutriate volume. Three ml were then pipetted into disposable cuvettes (5 reps $dilution^{-1}$. All cuvettes were incubated for 24 hours at 19°C at a light intensity or 4000 lux on a 12D:12L hour photoperiod.

Total mechanical stimulable light (bioluminescence) or TMSL was measured from each cuvette using the QwikLite/QwikSed test system. The system calculates mean bioluminescence (pmt counts in 30 seconds of stirring the cells), the standard deviation (pmt counts), the coefficient of variation (percent), and calculates light inhibition as a percent of control values for each dilution. An IC50 (an inhibition concentration of the elutriate that reduces bioluminescence by 50%) is normally calculated, although an IC25 was used to compare QwikSed toxicity results with amphipod survival and sea urchin development.

AMMONIA – CONFOUNDING EFFECTS

Because ammonia, associated with sediment, interferes or confounds QwikSed toxicity results, a protocol was developed to remove or reduce the impact ammonia would have on real toxicity associated with sediments. In the laboratory, total ammonia and pH was measured in 100% elutriates. From these data, the percent fraction or unionized ammonia was calculated for the 100% elutriates and then back calculated for all 5 dilutions. In the laboratory, we ran a series of ammonia standards and measured the percent light reduction. All ammonia levels were then converted to percent unionized ammonia, the most toxic portion of total ammonia. The observed reduction in percent of control light is assumed to be from ammonia. The observed difference is then added back onto the initial percent of control; any difference between this value and 100% (control) is then assumed to be from other toxicants, not ammonia. Two curves have been developed for both test species of dinoflagellates, *Gonyaulax polyedra* and *Ceratocorys horrida*. The curves give us a basic assessment as to which species may be impacted more by the presence of ammonia found in the sediment elutriates.

During the course of demonstrating the utility of using QwikSed as a fast screening tool at NAS Alameda and Pearl Harbor, other standard bioassay tests were employed to demonstrate a "standard response." Where possible, the same samples were tested by either all bioassay tests (amphipod, sea urchin development test, and QwikSed) or just 2 of the three tests.

Early in the demonstration, we wanted to compare the response of QwikSed to the 3-day sea urchin development test. An independent assessment was conducted by SAIC Environmental Testing Center in Narragansett, Rhode Island. Personnel were fully trained to operate a QwikLite toxicity unit and routinely conducted sea urchin (*Arabacia punctulata*) development tests. This assay has been used in regulatory programs in California and Washington state to assess the suitability of sediments (as elutriates) for ocean disposal activities. Other regulatory applications include usage of the test to meet minimum data requirements for the derivation of the U.S. EPA's Marine Water Quality Criteria and effluent testing.

The endpoint evaluated for in the sea urchin test was abnormal (*i.e*. delayed) development of the larvae. This response was measured in each of 3 elutriate concentrations per station/sample. The use of multiple concentration series for both the sea urchin and QwikSed provides information which can be applied to several techniques and integrated into the ecological risk assessment methodology. The concentration series responses can be used to develop an effect concentration (EC), a point estimate of the concentration that would cause a given percent reduction (IC_{50}) in development.

QWIKSED – SEA URCHIN TEST PROTOCOLS

Stock cultures of *Gonyaulax polyedra* were maintained at SAIC in Erlenmeyer flasks containing Enriched Seawater Medium (ESM) at 19°C. Cultures were held on a 12:12 hour (day:night) light cycle under cool white fluorescent bulbs at approximately 4,000 lux. Flask cultures were split weekly. Optimal cell culture concentrations, ranging from 3,000-5,000 cells ml^{-1}, were maintained during the testing period. *Gonyaulax* used for testing were evaluated during a reference toxicant test with zinc.

Adult sea urchins were obtained from a commercial supplier. A 12v transformer was used to electrically stimulate spawning. The urchins were segregated by sex into 20-liter aquaria each

holding about 15 animals. The aquaria were aerated and biological filters were used to maintain water quality. The tanks were partially renewed with filtered seawater from lower Narragansett Bay, RI twice weekly. Temperature was maintained at 15 ± 3°C. Salinity was between 28 and 32 ppt. Sea urchins used for testing were evaluated during a reference toxicant test with sodium dodecyl sulfate (SDS). The IC50 (the median inhibition concentration, in this case the inhibited response was fertilization) was evaluated against a control chart, a running plot of the IC50s obtained from 20 of the most recent reference toxicant tests performed at SAIC with *Arabacia punctulata*.

Modified U.S. EPA procedures were used to perform the sea urchin development test. Briefly, four male urchins were placed in seawater in shallow bowls. Males were stimulated to release sperm by touching the shell for about 30 seconds with the steel electrodes of a 12 v transformer. Sperm were collected using a 1 ml disposable syringe fitted with an 18-gauge, blunt tipped needle. The sperm were held on ice and were used within 1 hr of release. Sperm were diluted with seawater to a concentration of 5 $x10^7$ sperm ml^{-1}.

Four female urchins were placed in seawater in shallow bowls. Females were stimulated to release eggs by touching the shell as described above. Eggs were collected and held at room temperature for up to two hours with aeration. The eggs were washed two times with seawater by gentle centrifugation (500xg) for two minutes in a conical centrifuge tube. The eggs were diluted with seawater to a concentration of 2,000 eggs ml^{-1} and were aerated until used. Sperm and egg suspensions were mixed to a final concentration of 1:2,500 egg:sperm ratio.

After 60 minutes, 1 mL of fertilized egg suspension was added to 10 mL of sample in each of three replicates and was incubated for 48 hours at 20 ±1°C. The test was terminated by adding 2 mL of preservative to each vial.

One mL of suspension from each of the three replicates was transferred to a Sedgwick-Rafter counting chamber. Embryos were examined using a compound microscope (100X). One hundred embryos were examined for normal (i.e., not delayed) development as indicated by the presence of the pluteus larva. Additional subsamples from random replicates were examined when data varied by more than 10%.

The number of normal pluteii larvae per concentration for each sample was then recorded.

5.2 Data Assessment.
The primary objectives of this demonstration were to evaluate each of the three screening technologies in the following areas: (1) screening instrument/method performance compared to conventional sampling and analytical methods; (2) data quality (PARCC Parameters); (3) logistical and economic resources necessary to operate the technologies; and (4) range of usefulness in which the technologies can be operated and integrated into a screening procedure that allows more efficient assessment of sediment sites. Secondary objectives for this demonstration were to evaluate the technologies for their reliability, ruggedness, and ease of operation.

5.2.1 FPXRF.
FPXRF results from analyses of wet samples were compared with results from FPXRF results from analyses of wet samples were compared with results from standard analyses for samples collected from NAS Alameda. The data were used to determine correlation coefficients between the different methods. Figure 9 shows Fe, Zn, Cu and Pb results from the two methods plotted against one another. Linear regressions were calculated for each of the comparisons. The coefficients of determination (R^2), slopes, intercepts, and number of samples (n) are presented in Table 10.

In order to determine how FPXRF compared to standard analyses, the results from these two methods were plotted against each other. Figure 9 shows the Fe results from the pre-demonstration and demonstration samples for which both FPXRF field results and confirmatory results were available (n = 26). A linear trend can be observed. FPXRF results correlated well with results from the standard methods ($R^2 = 0.86$). In terms of instrument response, by pooling data together from the two deployments, it can be observed that there is little deviation in instrument response between the two different deployments (Southshore Pier area, 07/97 and Seaplane Lagoon, 10/98) in which samples were collected from different areas. It can also be seen that the FPXRF underpredicts the results when compared to standard methods. As discussed previously, FPXRF results (reported as wet weight concentrations) should be lower than results from standard methods, in which samples are concentrated by removing water (and in which results are reported as dry weight concentrations). However, the resultant reduction in sensitivity is acknowledged as part of a rapid field screening method that is used to delineate the location and extent of metals contamination in sediments as opposed to providing absolute quantitation of metals in the sediments.

A good correlation was observed between Zn results from FPXRF and standard methods (Figure 9). A linear trend is observed ($R^2 = 0.71$). As with Fe, the FPXRF underpredicted the results when compared to standard methods. Again, the response of the FPXRF remained the same over the course of more than one year with only a minor adjustment to source count times to adjust for source decay.

Copper results were also plotted to compare FPXRF to standard methods (Figure 9). The inset of the graph shows all of the results (n=26). A linear trend is not observed ($R^2 = 0.18$), however, one sample is dominating the poor correlation (Cu = 1,200 mg/kg by standard method). The preponderance of data falls in a lower concentration range. Therefore, a more meaningful statement about the instrument's linear response can be made by examining the data that fall in this range. The blowup of the graph shows the lower concentration range of the plot. A much-

improved linear trend can be observed, and the FPXRF correlates well with standard methods ($R^2 = 0.87$). However, it should be pointed out that the preponderance of the samples from the demonstration and a large number from the pre-demonstration had Cu values which fell below the instrument detection limit of 88 mg/kg, affecting the correlation coefficients. As with Fe and Zn, the FPXRF underpredicted the results as compared to standard analysis.

Finally, the lead results from the FPXRF method and standard methods were plotted together (Figure 9). Again, a good correlation was observed between the two methods ($R^2 = 0.81$). As was observed with the other elements, the FPXRF underpredicted the results as compared to the standard method. Some scatter around the regression is observed, which is most likely due to sample heterogeneity. As stated above, samples collected from the Southshore Pier area (07/97) and near the outfalls in Seaplane Lagoon were highly heterogeneous and contained coarse-grained particles consisting of pieces of wood from pier pilings, shell hash and other debris. As discussed in Section 5.1, because FPXRF is a bulk surface analytical technique, it is more sensitive to small-scale sample heterogeneity than a standard digestion analysis. However, all methods will be impacted by extensive sample heterogeneity.

In general, the FPXRF results compared well to results from standard methods. Good linear relationships were observed between FPXRF measurements and measurements from standard methods ($R^2 > 0.7$). In spite of an underprediction of results by FPXRF, the ability to rapidly delineate contaminant trends was demonstrated (see Section 5.3 for further discussion).

Comparison between On-site and In-laboratory FPXRF Results (NAS Alameda). As stated above, during the demonstration, 29 sediment samples were collected. During the first two days of deployment, 11 samples were collected from Seaplane Lagoon and were analyzed on site. The remaining 18 samples were collected the following week, and were subsequently analyzed in the laboratory. The 11 samples which were analyzed in the field were also re-analyzed in the laboratory. All 29 samples were sent out for confirmatory analyses as required by the regulatory project. Because the purpose of this report is to validate on site screening tools, the bulk of the data discussion pertains to results from the analyses performed on site (pre-demonstration and demonstration). However, a comparison of on-site and in-laboratory results allows for some insight into the repeatability of this method, as well as the effects of storing samples rather than analyzing them immediately.

Table 11 lists the on-site and in-laboratory results for Fe, as well as the on-site, in-laboratory and certified results for Cu, Pb and Zn. For both methods, the detection limits (DL) are listed as well. A number of things should be noted in these data. First, the 19 samples collected in regions outside Seaplane Lagoon, and analyzed only in the laboratory had, in general, lower Cu, Pb and Zn concentrations than did those in Seaplane Lagoon, and in many cases, levels below the FPXRF detection limit. These other samples had been selected primarily for their PAH content and potential toxicity, not for their metals content, while the Seaplane Lagoon samples were known to be rich in metals and organic contaminants. Second, the preponderance of the Seaplane Lagoon (SL) samples had certified Cu values below the FPXRF detection limit. Thus, while these data can be analyzed, they should be examined with caution, as the proportion of samples, which are appropriate for the instrument's capabilities, are limited.

With these caveats, Figure 10 compares the on-site and in-laboratory FPXRF results for Fe, Zn, Cu and Pb, respectively. Figure 11 shows results from both FPXRF measurements against results from certified analyses for Zn, Cu and Pb, respectively. As can be seen, Pb and Fe results for the two sets of FPXRF measurements correlate strongly. Both on-site and in-laboratory Pb FPXRF measurement correlate well with certified values, though the on-site measurement had a stronger correlation. The sample which can account for most of the difference between the on-site and in-laboratory measurements is SL06. This sample is quite different from the other samples in terms of Fe content (Table 11). This sample also had a significantly lower fines content (7% by weight clay and silt, as opposed to 46% for SL01, and 91-99% for SL02-SL05 and SL07-SL11, according to the contract analytical results). Thus, this coarse-grained sample would be very heterogeneous, and metals results will be very sensitive to sample mixing, differential settling, and subsampling effects.

There is some scatter in a few of the Zn results, resulting in lower correlation coefficients than are observed for Pb. This is most likely the result of some sample heterogeneity, as discussed above, but it is of note that the on-site results in this case correlated much more strongly with certified values than did the in-laboratory results.

As discussed, since Cu values were for the most part, lower than FPXRF detection levels, little can be said about the Cu results. However, it should be noted that while most Cu values in the on-site measurements registered as non-detects, the in-laboratory measurements reflected detectable (though still below the reliable detection level) levels of Cu. This is most likely due to the effect of sample settling during shipping and storage – some interstitial water is lost, essentially concentrating samples, and raising Cu concentrations to levels that are slightly detectable. Still, correlations with certified methods are poor, as is expected at these low levels.

Results from FPXRF analyses of wet samples were compared with results from standard analyses for samples collected from Pearl Harbor Naval Complex. Because of on-site instrument problems, for this demonstration, FPXRF results reported and compared to confirmatory analyses are based upon analyses made in the laboratory as opposed to in the field.

Figure 12 shows Fe, Zn, Cu and Pb results, as measured by FPXRF and by certified analyses. Linear regressions were calculated for each of the datasets. The coefficients of determination (R^2), slopes, intercepts, and number of samples (n) are presented in Table 12.

In order to determine how FPXRF compared to standard analyses at this site, the results from these two methods were plotted together. Figure 12 shows the Fe results from the pre-demonstration and demonstration samples for which both FPXRF laboratory results and confirmatory results were available (n = 65). A linear correlation was observed for Fe (R^2 = 0.7). As discussed previously, the mineralogical composition of the sediment in Pearl Harbor is quite different from that of west coast Continental US sediments (*e.g.* NAS Alameda). The sediments are composed of basaltic rock (iron-rich) and contain coral/shell hash (calcium carbonate) as opposed to the granitic rocks common to the west coast Continental US (primarily silica). The concentration of iron in sediments collected from Pearl Harbor ranges from less than 1% to 15% (as measured by standard methods). From the comparison of the results, it appears that the data cluster more tightly together in the lower concentration range (1-5%). Above this

concentration, the relationship between the two methods becomes less evident as more scatter is observed.

As discussed in Section 5.1, based on the observations made by comparing the results of FPXRF analyses to results from standard methods for the pre-deployment samples, it was expected that severe matrix effects would be encountered due to spectral interferences from the extremely elevated levels of iron. Physical matrix effects such as those caused by moisture and sample heterogeneity were also expected, although not to the degree encountered. Figure 12 also shows the results for Pb. A linear correlation was observed for Pb ($R^2 = 0.64$). The correlation was affected by the inability of the FPXRF to detect Pb at concentrations below ~ 50 - 75 mg/kg. A preponderance of the samples (64%) had Pb concentrations below 75 mg/kg, as determined by standard methods. Extreme heterogeneity in the samples collected from the Bishop Point site affected the comparisons of results between analytical methods as well. The samples collected contained large pieces of shell hash and metal debris, and thus the wet samples were difficult to homogenize. Such heterogeneity will affect both field screening and standard methods, reducing the probability that splits taken for the analyses have comparable concentrations, and resulting in poor correlations between analyses.

Copper results were also plotted, to compare results from FPXRF and standard methods (Figure 12). The inset of the graph shows all of the results (n = 65). A linear trend is observed ($R^2 = 0.71$), however one exceptionally high Cu sample is essentially controlling the correlation (Cu = 1,889 mg/kg by standard method). The preponderance of data falls in a lower concentration range (0 - 750 mg/kg). Therefore, a more meaningful statement about the instrument's linear response for these samples can be made by examining the data that fall in this range. The blowup of the graph shows the lower concentration range of the plot. A much weaker correlation is observed, ($R^2 = 0.46$). Several factors could be responsible for the reduced correlation, including chemical matrix effects (absorption of Cu x-rays by Fe), dilution effects caused by moisture content, sample heterogeneity (BP samples are known to have Cu-rich particles randomly distributed in sediments), and low concentrations of Cu (as measured by standard methods) present in these samples. Many of the samples (28 of 64 samples, or 44%, based upon the certified values) contain levels below the FPXRF method detection limit for Cu.

Finally, the results for Zn were plotted in the same manner (Figure 12). The inset of the graph shows all of the results (n=65). A poor linear correlation is observed ($R^2 = 0.46$), however two samples dominate this plot (Zn = 1,330 mg/kg and 1,570 mg/kg by standard method). The preponderance of data falls in a lower concentration range (0 – 1000 mg/kg). Again, by examining the data that fall in this range (blowup of the graph), a stronger correlation is observed ($R^2 = 0.59$).

Comparison of Results between Demonstration #1 and Demonstration #2.
Prior to this project, the FPXRF was evaluated as a technique for field analysis of metals in marine sediments. Sediment samples collected from many locations were analyzed and the results were compared to results from standard methods. Results from the evaluation suggested that, for most sediments, there is little variation in instrument response between the different sediments that were analyzed, suggesting that FPXRF is not very sensitive to sediment type (Kirtay *et al*, 1997, 1998).

Sediments from one of the demonstration sites used for this project were very different than sediments previously analyzed in terms of mineralogy, heterogeneity and source of contamination. Instrument response for Fe, Zn, Cu and Pb at NAS Alameda (Site #1) was consistent with results obtained from previous studies. However, the results obtained with sediments collected from Pearl Harbor suggest that FPXRF can be more sensitive to sediment mineralogy and heterogeneity than was previously thought. The high levels of Fe affected the ability of the instrument to detect Cu and possibly Zn. Because a sample preparation technique such as drying and grinding the samples prior to analysis by FPXRF was not employed during these demonstrations, sample heterogeneity also affected comparison of results between the FPXRF and standard methods.

Some of the problems encountered during this demonstration can be overcome by modifying the method used. Accuracy and precision can be improved by processing the samples (e.g., dry and grind) prior to analysis (Stallard *et al.*, 1995; Kirtay *et al.*, 1997). Field studies have shown that sample heterogeneity can have the largest impact on comparability with confirmatory samples (USEPA, 1998). Additionally, different calibration techniques can be used to improve FPXRF performance relative to standard analytical methods. For this project, a manufacturer-supplied method was used. This method is based on an "Effective Energy FP Calibration" routine. In essence, this calibration technique relies on pure element standards for FP calibration. The effective energy routine relies on the spectrometer response to pure elements and FP iterative algorithms to compensate for various matrix effects. Modifications such as an adjustment to y-intercept and slope of calibration curve based on instrument response to calibration check sample can be used. However, this technique was not employed for these demonstrations because the purpose was to evaluate the efficacy of a simple, universal method for screening sediment samples. Also, empirical calibration techniques using site-specific calibration standards can be used. All of these techniques should be considered when using FPXRF as a field screening tool. EPA Method 6200 (1998) provides detailed information regarding interferences and calibration techniques that can be used.

However, it should be pointed out that these demonstrations were carried out at particularly challenging sites – both with metal levels near or below the detection limit for some analytes, and one with a distinctive sediment type (from sediments adjacent to a volcanic island). In spite of both these factors, for samples with above the instrument detection limits, the FPXRF succeeded in its primary purpose: delineating contaminated from uncontaminated sediments, and allowing for the relative ranking of those sediments. This can be achieved without site-specific calibration or sample preparation to guide sampling and help delineate sites of concern. A second layer of analysis, either using site-specific calibration or sample drying and grinding, can aid in enhancing the accuracy of information gathered by FPXRF. This is more labor-intensive, but more rapid and cost-effective than certified analysis, and can be carried out on a large number of samples to interpolate between points characterized by the most costly certified analyses.

In order to carry out the goals of this demonstration/validation, data generated by FPXRF, UVF, and QwikSed technologies were compared to data generated by standard laboratory methods. However, it should be pointed out that, by definition, the screening technologies are semi-quantitative, and subject to different criteria than are the confirmatory analyses. Data quality parameters can be characterized by five indicators of data quality referred to as the PARCC

parameters: precision, accuracy, representativeness, completeness, and comparability. High quality, well-documented confirmatory laboratory results are essential for meeting the purpose and objectives of this demonstration. Therefore, the PARCC parameters, which can be used as indicators of data quality, were evaluated, where available, to determine the quality of data generated by the confirmatory laboratories. In addition, as appropriate, the PARCC parameters were utilized to evaluate the quality of data generated by each of the screening technologies. For the metals analyses by FPXRF, the following definitions were used for each PARCC parameter.

PRECISION

Precision refers to the degree of mutual agreement among individual measurements and provides an estimate of random error. Traditionally, precision of a technology is assessed with the use of field duplicate samples and the analysis of replicate sample measurements. Field duplicate samples provide precision data for sample collection, field preparation, handling, and transportation procedures. Replicate sample measurements provide data for the analytical precision of the specific technology. However, since FPXRF is carried out on bulk samples, without extraction, it is particularly subject to issues of sample heterogeneity (see previous discussions). To address this, an extra layer of sub-sampling was carried out on randomly selected field samples undergoing FPXRF analysis (see below).

Field duplicate samples were collected according to the sampling plan and analyzed. In this case, precision was evaluated in terms of the relative percent difference (RPD) between the results for these samples. During the on site demonstration at NAS Alameda (Demo #1) one set of field duplicate samples was collected. Results for the duplicate field samples (SL08 and SL11) are presented for both methods (Table 13). However, it should be noted that the field duplicate sample collected (SL11) is below the Cu detection limit, based upon certified analyses, so examining the RPD for the FPXRF values is of limited utility. During the demonstration at Pearl Harbor (Demo #2), two sets of field duplicate samples were collected from Bishop Point. There was some lack of precision between results for Zn and Pb as measured by FPXRF. Again, these results indicate the sediment was heterogeneous. Only a subset of the samples collected at Pearl Harbor was analyzed by standard methods and unfortunately, the field duplicate samples were not measured. Therefore, an evaluation of this parameter cannot be made for the standard methods used.

To evaluate instrument precision, triplicate measurements of randomly selected field samples were carried out. Instrument precision was evaluated in terms of the percent relative standard deviation (%RSD) between the replicate measurements and reported as Instrument %RSD. The percent RSD is defined as the standard deviation divided by the mean concentration times 100. Table 14 shows the results from the instrument precision analyses. When these results are plotted, the effect of concentration on precision can be observed (Figure 13). As the concentration of the target analyte increases, the precision increases. At low concentrations, instrument noise causes a signal which is a large proportion of the actual signal. As the analyte concentrations increase, noise is drowned out by analyte signal. This effect is most clearly seen with the results for Zn and Pb. As the concentration of these metals exceed two to three times the detection limit, the precision improves dramatically (< 20%). This effect is less apparent for copper, however, because the concentrations of Cu in these samples are too close to the detection limit to be relatively unaffected by noise.

To evaluate the issue of heterogeneity, randomly selected field samples were subsampled in triplicate and each subsample was analyzed by FPXRF. In this case, precision was evaluated in terms of the standard deviation (SD) and the %RSD between the replicate measurements and was reported as the heterogeneity-dependent RSD. The results for this evaluation are shown in Table 15.

If samples are highly heterogeneous, all analytical methods will be affected by this sort of heterogeneity – Table 15 and Figure 14 reflect the variability which comes from non-representative subsamples taken from a heterogeneous sample. These samples were collected from areas with very coarse-grained sediment that contained pieces of shell hash and other particles (e.g., paint chips, blasting grit). With an added sample homogenization method (*i.e.* dry and grind), disparate results from replicate FPXRF or other analyses of wet, heterogeneous sediments are unusual. Therefore, if this rapid screening technique is used, it is important to analyze multiple subsamples in order to account for this variability. For standard methods at heterogeneous sites, either multiple samples or the complete homogenization of a large sample is important as well.

Accuracy refers to the difference between a sample result and the reference or true value for the sample (Keith, 1991). It should again be stated that FPXRF is a field screening tool, which analyzes a thin layer of a bulk sample, and reports concentrations as wet weight, while most standard methods analyze extracts of samples, and report concentrations as dry weight. Thus, as discussed in previous sections, there are some problems inherent in comparing the results of these two approaches. However, in order to assess the accuracy of this screening tool, the assumption will be made that results from standard laboratory methods represent "true values", and FPXRF accuracy can be evaluated by comparison with these results.

A selected number of samples were sent out to analytical laboratories for standard confirmatory analysis for each of the demonstrations. The results from these analyses were compared with corresponding FPXRF results from the same samples. These corresponding data sets were subjected to linear regression analysis. The resultant R^2, slope, and intercept are used to characterize the accuracy of the FPXRF results, given the assumption that the standard method results represent "true values." In such a case, an R^2 and slope of 1 and an intercept of 0 would represent "perfect" accuracy. Previous field validations (Kirtay *et al.* 1997, 1998) of the Spectrace 9000 FPXRF used for this project suggest that an R^2 of 0.7 or greater, and a slope of about 0.5 can be expected, for the elements of interest. The results for each demonstration were discussed above.

Another parameter used to determine method accuracy and performance stability is the use of Performance Evaluation (PE) samples. Standard Reference Materials (SRMs) were analyzed with each set of demonstration samples. While these sediments are finely ground, dried, homogenized, and thus not truly representative of the sediments being analyzed at the site, they can still be used to provide an ongoing check and validation of instrument performance stability. Three SRMs from the National Institute of Standards (NIST #2704, #2710 and #2711) and marine sediment SRM (PACS-1) from the Canadian Research Council were used. The concentrations of each of the elements of concern ranged from the low parts per million (ppm) to thousands of parts per million. Although these samples are not representative of the samples analyzed for this project, the results indicate that the FPXRF underpredicts the results if directly

compared to the certified values, even in dry samples. These results are not unexpected. The application used for this project ("coarse-grain") is designed to accommodate grain size distributions; therefore a decrease in response is expected for some element concentrations with finely ground materials (TN Technologies, pers. comm.). The FPXRF proved to be consistent in terms of response for Fe, Zn, Pb and Cu as indicated by the small standard deviation between replicate measurements (Table 16).

REPRESENTATIVENESS

Representativeness refers to the degree to which the data accurately and precisely represent the conditions or characteristics of the parameter represented by the data.

As described in Section 5.3.1, FPXRF measures total elemental composition. Since total concentration of an element (*e.g.* Cu, Zn or Pb) is the parameter of interest FPXRF results are directly representative of the parameter of interest.

COMPLETENESS

Completeness refers to the amount of data collected from a measurement process compared to the amount that was expected to be obtained. For this demonstration, completeness refers to the proportion of valid acceptable data generated using each method. The completeness objective for data generated during this project is 95%.

For the goals of this project, 100% of the data generated from both methods were considered valid, acceptable data for Demonstration #1. For Demonstration #2, none of the data generated on site by FPXRF was considered valid due to instrument performance problems. Therefore, the samples were reanalyzed in the laboratory by FPXRF subsequent to instrument repair. One hundred percent of the laboratory FPXRF-analyzed samples and 100% of the data as generated by standard methods were considered valid and acceptable.

COMPARABILITY

Comparability refers to the confidence with which one data set can be compared to another. Comparability of the FPXRF results was evaluated by an examination of the precision, accuracy and instrument stability as described in sections above.

Logistical and Economic Resources Required. Examples of the logistical and economic resources required for the deployment of the FPXRF to a demonstration site is provided in Section 6 (Cost Assessment). Also included is an example of the cost per sample for laboratory screening by FPXRF.

Range of Usefulness for Efficient Assessment of Sediment Sites. Contaminant distribution maps using field screening data can be rapidly generated on site to provide a visual image of the relative levels of metals contamination of the area(s) under survey. At NAS Alameda (Demo #1), the FPXRF results from the on site rapid analyses were combined with spatial data (sampling coordinates) to create maps. A comparison of the maps produced with on site data to maps produced several months later using the certified data show similar patterns of contamination for Zn and Pb (Figure 15). Both the corners of Seaplane Lagoon as well as the

inboard pier area were identified as areas of higher contamination relative to the other area surveyed. Such a comparison shows that the FPXRF can be used at this site to rapidly delineate areas of concern in order to select samples for subsequent detailed analysis using more expensive, quantitative analytical methods. A similar type of comparison was made for the second demonstration site, Pearl Harbor. In this case, the results for Cu, Zn and Pb from the laboratory FPXRF measurements were mapped and compared to the maps generated using certified data. In terms of comparability based on linear regression analysis the FPXRF did not compare well to standard methods. However, if the screening data are used to identify contaminant trends, the FPXRF was able to flag certain regions (*e.g.* South East Loch) that contain higher levels of Cu and Zn relative to other areas of the Naval Complex (Figure 16 and Figure 17). Due to sample heterogeneity issues, the power of discrimination for Pb by FPXRF was less evident at this site (Figure 18).

5.2.2 UVF.

The UVF screening data were compared with results from standard lab analyses in Table 17 through Table 19. Figure 19 through Figure 21 show results from the two methods plotted against one another. The data were used to determine correlation coefficients (R^2) between the different methods. The correlation coefficients in the three figures range from 0.71 to 0.89. As discussed in the previous sections, this variation in slope is between sites related to the change in composition of the mixture of PAHs in each area. Seaplane Lagoon shows the lowest slope and represents an area with a historic PAH source and weathered signature. The greater proportion of heavier and alkylated PAHs compared to lighter and parent PAHs leads to higher fluorescence at any given total PAH level. The Pearl Harbor and Alameda Pier samples show a fresher creosote source PAH signature with greater proportions of lighter and parent compounds which requires higher total PAH levels to reach the same fluorescence intensity. This supports the requirement that some percentage of site samples should always be sent for lab confirmation to generate a site specific calibration curve to generate absolute PAH levels. UVF data without lab calibration data can provide relative ranking of PAH levels within a site with similar PAH compositions, but without knowledge of the mixture of PAHs present it would be difficult to pick an appropriate calibration curve to convert UVF data into PAH levels.

For this project, multiple field events occurred at NAS Alameda and Pearl Harbor where over 100 samples were screened for PAHs. Confirmatory analyses were carried out on a subset of samples. The results from these samples with both screening and confirmatory results are provided in Table 17 through Table 19. In these tables the sample numbers are followed by the dry weight corrected fluorescence values measured in intensity units. This corrected value is simply the raw instrument intensity values divided by the dry weight proportion of the original wet sample. This corrects for the variation in moisture content (and therefore solid matrix) due to the range in grain sizes in the measured samples. The next two columns in these tables are standard deviations and percent relative standard deviations for a subset of the samples where multiple sample splits were run through the entire method (both extraction and analysis steps). The next column in the tables is the dry weight corrected lab confirmation values run by a commercial analytical chemistry lab (AD Little Inc., Duxbury MA) using standard EPA method 8270, gas chromatography with mass detector (GC/MS) modified for selective ion monitoring to achieve lower detection limits (2 ng/g for individual PAHs). These total PAH values are the sum of 16 individual parent and some related alkylated PAHs. In addition to the 16 EPA priority pollutant PAHs (all parent compounds), additional PAHs (including some alkylated forms) were measured because many different PAHs contribute to the fluorescence signature of the samples. Even this extended list of PAHs doesn't cover all the possible PAHs that may contribute to the fluorescence, but it is enough to see some of the major parent and alkylated contributors.

The results in Table 17 include data from the deep water piers outside of Seaplane Lagoon at NAS Alameda collected during pre-demonstration events in 1997, along with the actual demonstration in October of 1998. Table 18 includes data from Seaplane Lagoon at NAS Alameda collected during the pre-demonstration event in December 1997 and the actual demonstration in October of 1998. Table 19 contains data from Pearl Harbor collected during the demonstration in February of 1999. Although additional pre-demonstration samples were available from Pearl Harbor from an ongoing regulatory project, data are not reported here because sampling for fluorescence occurred from archived samples 2 years after initial sampling (and outside the normal 2 week holding time for extractions to begin). These Pearl Harbor pre-

demonstration samples were used to determine range and any potential problems with running Pearl Harbor sediments. The majority of the reported samples were extracted within 24 hours of field collection, with sample extracts run on the fluorometer within a few days of collection. Due to a delay in sampling for the regulatory project at Alameda, some samples collected in October of 1998 were sent back to the lab and extracted within a week of collection, with fluorescence measured within 2 weeks of collection. Hexane extracts run over a period of several weeks show little variation in fluorescence as long as vials are capped tightly and evaporation is avoided. Unlike the XRF results, UVF values (for the same extract) of samples do not vary from initial field measurement to later (several weeks) laboratory re-measurements. Since UVF data don't seem to show variations from these minor differences in method, all data will be discussed together regardless of whether it was run immediately in the field or several weeks later in the lab. This is similar to conventions used in standard lab procedures UVF data will be compared with, where the lab confirmatory data may have up to 2 weeks between the time samples are extracted from the time of collection.

As appropriate, the PARCC parameters were utilized to evaluate the quality of data generated by UVF screening technology.

PRECISION

Precision refers to the degree of mutual agreement among individual measurements and provides an estimate of random error. Traditionally, precision of a technology is assessed with the use of field duplicate samples and the analysis of replicate sample measurements. Field duplicate samples provide precision data for sample collection, field preparation, handling, and transportation procedures. Replicate sample measurements provide data for the analytical precision of the measurement method. For UVF the instrument precision value for replicate measurements of the same hexane extracts is below 1% RSD. The majority of the variation in the RSD values in Table 17 through Table 19 is due to sample heterogeneity and extraction variability. Most samples are below 10% RSD, with 1 out of 14 samples run in replicate samples in the 20% RSD range. These samples with higher %RSD levels tend to be at the higher PAH range, where PAH levels tend to be more heterogeneously distributed.

ACCURACY

Accuracy assesses how close the screened value is to the true value for the sample. In order to assess the accuracy of UVF the assumption will be made that results from standard laboratory methods represent "true values", and UVF accuracy can be evaluated by comparison with these results. Figure 19 through Figure 21 show these relationships, with correlation coefficients in the range from 0.71 to 0.89. A perfect correlation between screening and lab data would result in a value of 1.0, so these relationships are considered good. As mentioned earlier, sample heterogeneity can result in sample splits showing variations, so some variability in the relationships between screening and lab results is expected related to these problems.

REPRESENTATIVENESS

Representativeness refers to the degree to which the data accurately and precisely represent the conditions or characteristics of the parameter represented by the data. UVF does not measure

PAH concentrations directly, but a bulk fluorescence that can be directly related to PAH concentrations. The ability of this fluorescence to represent the actual PAH concentrations is subject to the limitations discussed previously.

COMPLETENESS

Completeness refers to the amount of data collected from a measurement process compared to the amount that was expected to be obtained. For this demonstration, completeness refers to the proportion of valid acceptable data generated using each method. The completeness objective for data generated during this project is 95%. For the goals of this project, 100% of the data generated were considered valid, acceptable data. One hundred percent of the laboratory data generated by standard methods were considered valid and acceptable.

COMPARABILITY

Comparability refers to the confidence with which one data set can be compared to another. Comparability of the UVF results was evaluated by an examination of the precision, accuracy and instrument stability as described in sections above.

Logistical and Economic Resources Required. Examples of the logistical and economic resources required for the deployment of the UVF to a demonstration site is provided in Section 6 (Cost Assessment). Also included is an example of the cost per sample for laboratory screening by UVF.

Range of Usefulness for Efficient Assessment of Sediment Sites. Collection of data necessary to support decisions at Navy marine sites in a cost-effective manner is often hindered by the complexity and heterogeneity of marine ecosystems. Detailed site investigations require extensive sampling and subsequent laboratory analyses for both metal and organic contaminants. Samples are often collected without any *a priori* knowledge of the nature and extent of contamination. Due to the high cost of laboratory analyses, the number of samples taken is often cost-limited. Thus, zones of contamination can be missed, or, if located, over- or under-estimated. For more detailed spatial information on the extent of contamination, sites of interest must often be sampled and analyzed in an iterative manner. Chemical assays are often combined with additional laboratory analyses; including one or several bioassays to determine whether there are adverse biological effects of these contaminants in various media (*e.g*. sediment, elutriate, water column). This approach can be prohibitively costly, slow and labor-intensive. When used appropriately, rapid sediment characterization tools can streamline many aspects of the site assessment process, delineating areas of concern, filling information gaps and assuring that expensive, certified analyses have the highest possible impact.
Contaminant distribution maps using field screening data can be rapidly generated on site to provide a visual image of the relative levels of metals contamination of the area(s) under survey. At NAS Alameda, UVF results from the on-site rapid analyses were combined with spatial data (sampling coordinates) to create maps just like what was done for XRF data. Figures 22 and 23 show examples of how UVF data can be plotted to show spatial relationships.

5.2.3 QwikSed.

Twenty sediment samples were collected from Seaplane lagoon (SP), Pier Area 1 (PA), and Breakwater Beach (BB) at the Naval Air Station. The sediments were later shipped to SAIC for side-by-side comparison toxicity studies. At SAIC Narragansett, elutriates of test sediments were prepared for the dinoflagellate and sea urchin tests. Preparation began by adding homogenized sediment to filtered natural seawater collected from Narragansett Bay, RI on an incoming tide in a 1:4 volumetric ratio. The sediment slurry was handled as previously described. Dilutions were prepared by mixing the filtered supernatant with ESM or filtered natural seawater collected from lower Narragansett Bay on an incoming tide for the dinoflagellates and sea urchin tests, respectively. Elutriate dilutions (25%, 50%, and 100%) as well as an ESM performance control (0%) were tested.

The mean light output, standard deviation and coefficient of variation was calculated for each dilution and each sample for the dinoflagellate test. The mean number of normal larvae, the standard deviation, and the coefficient of variation was calculated for each dilution and each sample for the sea urchin test. For both tests, the calculated mean for each test concentration was compared with the control to normalize against the control response. A scatter plot representing the IC_{50}'s response for both QwikSed and the sea urchin development test (percent of control) on the same 20 sediment samples indicated a highly significant correlation ($r^2 = 0.943$). Interpretation of this relationship shows that both test organisms respond similarly to the toxicity, or lack of toxicity encountered in the sediment elutriates.

QwikSed (1 Day Test) vs the Amphipod (10 Day) Survival Test and Sea Urchin Development Test (3 Day), Alameda.

Twenty-five sediment samples were collected in Alameda and processed on-site (Table 20). Elutriates were prepared for the one day QwikSed toxicity test using *Gonyaulax*. Greatest toxicity (40-64% of control cells) with QwikSed was exhibited at the pier areas as was in the amphipod toxicity test. Less toxicity was observed in the lagoon area of NAS Alameda. Overall, both QwikSed and the amphipod tests observed toxicity at 5/25 of the same stations (20% of total samples) while both observed no toxicity at 18/25 stations (72% of the total) for an overall agreement of 92% of all samples (Table 21). QwikSed and the sea urchin development test displayed intermediate agreement of 76% of all samples. Only 2 /25 (8% of all samples) samples were both toxic to QwikSed and the sea urchin development test while 17 of 25 samples were not toxic in both tests (68% of all samples) for an overall agreement of 76% (Table 22). The worst relationship at NAS Alameda was between the amphipod survival test (10 day) and the sea urchin development test (3 day) (Table 23). Again, only 2 of 25 samples were toxic to both tests (8%) while 16 of 25 samples were not toxic in both tests (64%) with both in agreement for 72% of all samples.

QwikSed (1 Day Test) vs the Sea Urchin Development Test (3 Day), Pearl Harbor.

Eighteen sediment samples were collected from Middle Loch and Bishop Point. All sediments were immediately placed into a refrigerator until leachates could be prepared that day or the following day. Duplicate leachate samples were also collected and shipped overnight to MEC Analytical Systems, Inc. in Carlsbad for concurrent testing with the 3-day sea urchin development test. MEC laboratory personnel received all sediment leachate samples on 6

February and 12 February 1999. Control seawater was collected on 4 February from Scripps Institute of Oceanography and held in a re-circulating system until test initiation. Toxicity tests were conducted on 6 February to 9 February and from 12 February terminated on 15 February 1999. Toxicity tests were conducted on the purple sea urchin Strongylocentrotus purpuratus. All methods and procedures employed followed general guidelines established by the EPA in Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to West Coast Marine and Estuarine Organisms (EPA-600/R-95/136), August 1995, and Annual Book of ASTM Standards, Water and Environmental Technology, (ASTM, 1998).

Echinoderm larvae were exposed to concentrations of 0,6.25, 12.5, 25, 50, and 100%. No adjustments or manipulations were carried out on samples. Mean control development was 97% after 72 hours. The following samples had one or more concentrations exhibit a significant reduction in development: BP 4, BP 5, BP 6, BP 7, BP 8, BP 9, BP 10, BP 11, BP 12, BP 13, BP 14, BP 15, and BP 21.

In conjunction with the 3-day toxicity tests, a reference toxicant test, was also conducted to assess purple sea urchin sensitivity. The reference toxicant was copper sulfate, tested at nominal concentrations of 2.5, 5, 10, 20, and 40 ppb. The calculated 72-hour LC_{50} was 16.8 ppb, which falls within the acceptable reference toxicant range listed in the EPA manual (10.1 – 22.5 ppb). The calculated 72-hour LC50 for the first set of samples was 21.4 ppb.

Water quality parameters were within the recommended limits except for salinity. Salinity in the 100% leachates for BP 4, BP 5, BP 6, BP 10, BP 11, BP 12, BP 17, BP 21, ML 4, ML 12, ML 13, and also the 50% concentration for BP 10, BP 11, BP 12, BP 17, ML 4, ML 12, and ML 13, were between 28.7 and 31.0 ppt. This was just below the recommended test salinity of 34 ± 2 ppt. Purple urchins are found along the Pacific Coast from California to Washington. As intertidal creatures, they are able to handle sudden changes in their environment including salinity. The salinity range for echinoderm testing from "Recommended Guidelines for Conducting Laboratory Bioassays on Puget Sound Sediments (PSEP), July 1995" is 28±2 ppt. Locally collected urchins and sand dollars have been successfully tested at 28 ppt at the MEC laboratory. No salinity adjustments were made to the above mentioned samples. This minor deviation was not expected to have a negative impact on the final outcome of these results. Lastly, ammonia did not appear to have confounded test results as even sample ML4 with the highest ammonia level had 98.9% normal development.

Of the 18 samples, 13 samples proved to be toxic to both the QwikSed and the sea urchin development test while 2 of the 18 samples showed no toxicity in both tests (Table 25). Two samples were not toxic to QwikSed, but toxic in the sea urchin test, while 1 sample was toxic to QwikSed and not in the sea urchin test. Overall, 72% of the samples were toxic to both tests while 11% of the samples were not toxic in both tests for an overall agreement of 83% (Table 25).

5.3 Technology Comparison.

5.3.1 FPXRF.

There are other commercially available instruments (techniques) that can be used for the on-site analysis of environmental samples. However, these techniques are not necessarily suited to operation in a rugged field environment, or for the analysis of sediment samples (solid matrices). For example, Graphite Furnace Atomic Absorption (GFAA) Spectroscopy has been used primarily in the field for the analysis of metals in water. GFAA could be used to determine metals in soil or sediment, but the sample preparation for metals in these matrices is extensive and is not practical for field applications. GFAA cannot be described as a truly field portable instrument. GFAA instruments are extremely sensitive and therefore, must be operated in a clean, climate controlled environment. This can be difficult but not impossible to achieve in a field environment. In addition, the 220-volt electrical power requirement often precludes remote operation. However, GFAA is an example of taking the laboratory to the field. Miniaturization of electronics has significantly reduced instrument size and weight, making it easier to use the instrument in a field laboratory (*e.g.* mobile laboratory) (USEPA, 2000). On the other hand, Energy Dispersive X-ray Fluorescence (EDXRF) Spectrometry is a method that is used for detecting metals in soil and sediment. Some of the primary elements of environmental concern that EDXRF can identify are arsenic, barium, cadmium, chromium, copper, lead, mercury, selenium, silver, and zinc. Field-portable X-ray fluorescence (FPXRF) units that run on battery power and use a radioactive source were developed for use in analysis for lead-based paint and now are accepted as a stand-alone technique for lead analysis. In response to the growing need for field analysis of metals at hazardous waste sites, many of these FPXRF units have been adapted for use in the environmental field. The field-rugged units use analytical techniques that have been developed for analysis of numerous environmental contaminants in soils. They provide data in the field that can be used to identify and characterize contaminated sites and guide remedial work, among other applications (USEPA, 2000).

A list of the major advantages and disadvantages of these two technologies is provided in Table 26. Site-specific considerations must be considered in determining the appropriateness of using FPXRF, or any technology, for the rapid characterization of metals content in sediment samples. Data analysis and interpretation is likewise dependent on site-specific considerations as illustrated in this report.

5.3.2 UVF.

A number of additional field screening methods are available to screen for the presence of PAHs in sediments. Field deployable GC/MS units are one alternative, but like GFAA discussed above these typical laboratory techniques are more difficult to operate in the field. In comparison to UVF, field deployable GC/MS units are more labor intensive, costly to operate, and difficult to maintain under field conditions. Immunoassay test kits have in the past been less quantitative than desired, with results often being list as above or below a specific field PAH standard value. More recent immunoassay developments have led to more quantitative results, and these techniques can now be considered for field screening of PAHs.

5.3.3 QwikSed.

QwikSed, as a screening technology, compared favorably with the sea urchin larval development test in that the same trends were exhibited by both tests from sediments collected at NAS Alameda. The pier areas exhibited the greatest toxicity by both the sea urchin development test and QwikSed while no toxicity or very little toxicity was observed in Sea Plane Lagoon. While both organisms are sensitive to ammonia and the associated toxicity, they also appear to be similar to other toxic components associated with sediment leachates. The strength in the QwikSed test lies with a lesser amount of time required to conduct the test than the sea urchin or the amphipod test. An initial assessment of the problem areas can be determined with the 1 day turn around results from QwikSed versus the costs of shipping samples out of the study site to a contractor lab to conduct either the sea urchin development test or the amphipod survival test. Once the toxic areas of concern are identified with the screening technologies, then the more "standard" toxicity tests could be employed to address regulatory issues.

The sea urchin and QwikSed were in agreement for 76% and 83% of all samples tested at NAS Alameda and Pearl Harbor, Hawaii. Surprisingly, the pre-demonstration amphipod survival data and the demonstration QwikSed data at NAS Alameda showed the best agreement (92%) for any combination of comparisons among the toxicity tests. However, the screening user would save 85% of the costs of conducting the amphipod sediment tests by using QwikSed as a screening tool for identifying the toxic areas (Table 27).

5.3.4 Screening Technology Integrated Summary.

The previous sections of this chapter have validated the screening data by showing high correlations to standard laboratory measurements. This demonstrates each screening technique produces high quality data that can be used in a variety of study designs. This section will present case studies to show how these screening data can be integrated into various study designs to produce a cost effective, efficient study.

At the Alameda demo site, the screening tools were used to assist an ongoing regulatory project by providing information to assist in the positioning of their sample locations. During pre-demo sampling preliminary contour maps were generated to allow the regulatory sample positions to be selected. Regulatory samples were selected to span the observed chemical gradients at the site to ensure full range exposure-effects relationships would be generated. The Alameda demo was performed at the same time as the regulatory sampling to provide "real-time" feedback to ensure samples were actually collected over the full chemical gradients at the site. By performing the pre-demo screening prior to the actual regulatory sampling, the screening contour maps were available for discussions with the regulators to show where regulatory samples would be located. This allowed regulator input and promoted regulator acceptance of the sampling design for the regulatory project. Putting the screening tools in the field for the actual site demo at the same time the regulatory project was sampling provided regulators the opportunity to observe the screening techniques as well as providing the project with "real-time" feedback. The full range of laboratory measurements done for the regulatory project on sample splits also supplied validation data at no cost for the demo.

Future deployments of the screening techniques need only use portions of this procedure depending on their needs. For example, screening techniques could be used in the pre-demo role discussed above at sites with little available information on contaminant distributions. Figure 22 shows pre-demo PAH screening data from deep water pier areas at Alameda. A limited amount of demo data from this area (3 sites along the Quay Wall) showed elevated contaminant levels. This screening contour map shows the elevated levels of PAHs are restricted to outfall areas along the Quay Wall and allowed the regulatory sampling to focus in these areas. Or the screening tools could be deployed as they were during the actual Alameda demo, to serve as a "real-time" monitor to ensure full contaminant gradients were being sampled. Figures 23 - 26 show both screening and laboratory data from the Alameda demo, where measurements of a number of parameters that show wide gradients were observed. In the Seaplane Lagoon area, existing laboratory measurements were available to generate contour maps for discussions with the regulators. However, sample heterogeneity made it difficult to ensure the full contaminant gradients would be sampled to generate needed exposure-effects relationships with a limited number of regulatory samples. The screening techniques were used in a near real-time manner to ensure that the samples collected for laboratory analyses would cover the expected gradients.

The objectives and use of screening techniques for the second demo at Pearl Harbor were slightly different. At that site, the regulatory project had defined strata (sampling areas) from which a single sample was collected with the underlying assumption that measured values would be representative of that strata. The screening data were used to help define heterogeneity and decide how representative single sample values were for the strata. Figures 27 – 29 show some of these data from the demo at the Bishop Point location in Pearl Harbor. At Bishop Pt., the existing regulatory data show a single sample from between two piers in Stratum 2 with elevated

contaminant levels and depressed amphipod survival. The screening data indicate this regulatory sample is not representative of the whole stratum, but only the sediments right in front of the quay wall. By integrating the different screening results together it is possible to develop a "weight of evidence" scheme similar to those used with standard laboratory data to differentiate areas of concern. For example, Figure 29 shows several areas with multiple "hits" (as defined in the figure) from the several screening techniques. This type of procedure could be used to prioritize areas for additional work or discussions with the regulators about the existing regulatory data. It should be noted that the contouring of this screening data is very preliminary, and the heterogeneity suggests very different results are possible by varying initial assumptions. This supports the use of this type of screening data to assess area heterogeneity and to support discussions with regulators about how much data are required within a stratum to make a regulatory decision. For other areas of Pearl Harbor (for example Middle Loch), data are much more homogeneous and single samples within strata might provide sufficient confidence to reach a decision.

6. Cost Assessment

6.1 Cost Performance.

In addition to the technical performance of the screening techniques, the cost of screening technique use plays a major role in determining whether screening will prove to be a useful addition to sediment assessments. As a general rule, screening techniques are inexpensive when compared to traditional standard laboratory techniques. This advantage needs to be weighed against potential limitations discussed in Chapters 2, 4, and 5. To more fully evaluate the cost performance of the screening techniques, the following tables provide assumed costs for use under different scenarios. Due to the costs involved for mobilization and demobilization to deploy onsite, cost examples are given for both deploying onsite as well as having samples sent back to a centralized lab facility (similar to what would be done for standard analyses). Because there is economy of scale for most of these screening techniques, running larger sample sizes will generally prove more cost-effective. This is demonstrated in the examples with cost estimates for assessing both 30 and 100 samples at a time. Table 28 through Table 30 contains examples for XRF, UVF, and QwikSed, respectively. Table 31 contains estimates for annual operation by technology type and combined.6.2 Cost Comparisons to Conventional and Other Technologies.

Table 28, Table 29, and Table 30 contain the approximate per sample cost currently charged by laboratories for the standard laboratory analyses. For standard sediment metals analyses of Cu, Pb, and Zn, the cost would run between $150 to $300 depending on the laboratory. For PAHs the per sample cost is around $500, and would include a breakdown of the individual PAHs as well as the total given by the screening technique. For bioassays, the cost is highly variable depending on the particular bioassay. Sea Urchin larval development bioassay run around $500, while the amphipod bioassay may run up to $1500. For this comparison in the above table, an average of $1000 is used.

All three technologies are easily transferred and shippable. FPXRF, UVF and QwikSed materials can be transferred as luggage aboard commercial flights. Total weight of each is ranges between 150 and 300 lbs. Each technology is contained within a protective carrying case and does not need special handling requirements. Centrifuge, test chamber, and miscellaneous laboratory supplies may be more appropriate to ship ahead of time.

7. Regulatory Issues

This demonstration project was designed as part of an ongoing regulatory project to encourage interaction and involvement with regulators. By collecting field screening and standard laboratory data on the same samples during a regulatory project, acceptance of the screening tools will be promoted. During the regulatory process, public participation is allowed through meetings where project status and results are discussed.

7.1 Approach to Regulatory Compliance and Acceptance.
In addition to these ESTCP demonstrations, EPA has had screening techniques demonstrated in several programs, including the SITE and ETV programs. Due to the involvement of and interactions with regulators during all these demonstrations, screening techniques are becoming more accepted at sediment assessments in the same manner they are in soil sites. Additionally, EPA is including more screening methods in their standard SW-846 manual of accepted analytical techniques. The proof can be found at websites such as http://clu-in.org/char1.htm, where information and discussions about regulatory acceptance of innovative techniques such as screening are present.

8. Technology Implementation

8.1 DoD Need.

Within the Department of the Navy alone, there are an estimated 110 facilities with sediment contaminant sites with assessment needs which carry estimated costs of over $500 million (NAVFAC NORM database). These figures are expected to be even greater for the DoD as a whole. Given the assumptions in Chapter 6 on cost implementation, analytical costs could be expected to reduced by a conservative 50% if screening techniques were integrated into existing laboratory based assessment programs.

8.2 Transition.

The transition plan for screening techniques within the Navy is already in progress. The jointly funded ESTCP-NAVFAC demos reported in this report provide the basis for case studies to show screening utility. Additional case studies are available from other sites, including EPA SITE and ETV programs. A series of RITS (Remediation Innovative Technology Seminar) classes during October 2000 at 8 NAVFAC sites around the country were used to transition information to RPMs from Navy sites. NAVFAC has contracted Battelle (Columbus, OH) to run these classes and put together a screening guide for RPM use. This guide will provide RPMs with a short review of screening techniques, SOPs for various screening techniques, and SOWs (statement of work) with generic contract language to facilitate screening use at Navy sites. Since most environmental work at Navy sites is performed by contractor, the transition of these screening techniques mostly occurs via contractor utilization. RPMs must be given the authority to allow screening technologies to be employed by DoN contractors, including policy that screening should be included as needed in an efficient, cost-effective assessment.

9. Lessons Learned

Much of the cost in demonstrations of innovative technologies is in analytical laboratory costs. Each of these screening techniques required laboratory validation data as part of the demo. By partnering with ongoing NAVFAC regulatory projects, many of these laboratory costs were paid by the regulatory project since these laboratory measurements were a required element of their project. Unfortunately the timetable for the ESTCP project then becomes dependent on the regulatory project, which is often delayed for numerous reasons.

Although the regulator community was initially suspicious of screening techniques (due to concerns of adequate detection limits, matrix effects, fear of replacing all laboratory data, etc.), once their concerns were addressed they actually became strong advocates of screening techniques. If any innovative techniques are to be successfully employed, successfully addressing regulator concerns are an important component of the process.

10. References

Air Traffic Control NAS Alameda. 1992. "International Station Meterological Climate Summary - 1950 to 1985, Year/Month Total Precipitation (Inches) from Daily Observations," Division OPS, Building 19, NAS Alameda, California.

American Society for Testing and Materials, "*Annual Book of Standards, Water and Environmental Technology*, Volume 11.05," Biological Rate and Environmental Effects. American Society for Testing and Materials, Philadelphia, PA., 1995.

Apitz, S E., Borbridge, L M., Theriault, G. A., Lieberman, S. H., "Remote in situ determination of fuel products in soils: field results and laboratory investigations," *Analysis,* Volume 20. p. 461-474, 1992.

Bertin, E.P., "Principles and Practice of X-ray Spectrometric Analysis," Plenum Press, N.Y., p.1079, 1975.

Bulich, A.A., "Use of luminescent bacteria for determining toxicity in aquatic environments," in (L.L. Markings, R.A. Kimerle, eds.) Aquatic Toxicology, Philadelphia: American Society for Testing and Materials, p 98, 1979.

Carlson, P. and D. McCulloch, "Aerial Observations of Suspended-Sediment Plumes in San Francisco Bay and The Adjacent Pacific Ocean," Journal Research U.S. Geologic Survey. Volume 2, Number 5, September - October, 1974.

Chapman, P.M., R.N. Dexter, and E.R. Long, "A Field Trial of the Sediment Quality Triad in San Francisco Bay," National Oceanic and Atmospheric Administration (NOAA) Technical Memorandum NOS OMA 25, 1986.

COE, "Wetlands Delineation Manual," Technical Report No. Y-87-1, U.S. Army Engineering Waterways Experiment Station, Vicksburg, MS, 1987.

COE, "Sediment Budget Study for San Francisco Bay, Final Report," Prepared by Ogden Beeman & Associates, Inc., Portland, Oregon. Prepared for San Francisco District, Corps of Engineers. February, 1992.

COE and U.S. Environmental Protection Agency (USEPA), "Testing Manual for Evaluation of Dredged Material Proposed for Ocean Disposal," USEPA 503/8-91/001, 1991.

Corbitt, R.A.,"*Standard Handbook of Environmental Engineering*," McGraw-Hill. New York, 1989.

Ecology and Environment, Inc. (E&E), "Initial Assessment Study, Naval Air Station, Alameda, California," Prepared for Navy Assessment and Control of Installation Pollutants Department, Naval Energy and Environmental Support Activity, Port Hueneme, California, 1983.

Filkins, J., "Draft experimental methods: Estimates of PAHs in lacustrine sediment by fluorometry," USEPA-Large Lakes Research Station, unpublished interim report, 1992.

Gilbert, R.O., "Statistical methods for environmental pollution monitoring," Van Nostrand Reinhold Co. In., New York, p.320, 1987.

Harding Lawson Associates (HLA), "Closure and Post Closure Maintenance Plan. Alameda Naval Air Station Landfill," Alameda Naval Air Station, Alameda, California. October 28, 1988.

Hargrave, B.T., and G.A. Phillips, "Estimates of oil in aquatic sediments by fluorescence spectroscopy," Environ. Poll., V.8:193-215, 1975.

Kieth, L.H., "Environmental Sampling and Analysis - A Practical Guide," Lewis, 1991.

Kirtay, V.J, J.H. Kellum, and S.E. Apitz, "The Validation of Field-portable X-ray Fluorescence Spectrometry for the Analysis of Metals in Marine Sediments," Proceedings from an international symposium on field screening methods for hazardous wastes and toxic chemicals, vip-71, 822 - 837, 1997.

Kirtay, V.J, Kellum, J.H. and Apitz, S.E, "Field-portable X-ray fluorescence spectrometry for metals in marine sediments: results from multiple sites," Water Science and Technology (in press), 1998.

Lapota, D., D.E. Rosenberger, and D. Duckworth, "A bioluminescent dinoflagellate assay for detecting toxicity in coastal waters," in Bioluminescence and Chemiluminescence, Fundamentals and Applied Aspects (A.K. Campbell, L.J. Kricka, and P.E. Stanley, eds.), John Wiley & Sons, Ltd., Chichester, England, 1994, pp.156-159.

Lapota, D., D. Duckworth, and D. Rosenberger, "Proposed standard guide for conducting toxicity tests with bioluminescent dinoflagellates," Draft ASTM procedure, 1997.

Long, E.R., D.D. MacDonald, M.B. Matta, K. VanNess, M. Buchman, and H. Harris, "Status and Trends in Concentrations of Contaminants and Measures of Biological Stress in San Francisco Bay," NOAA Technical Memorandum NOS OMA 41, 1988.

Long, E.R., D.D. MacDonald, S.L. Smith, and F.D. Calder., "Incidence of Adverse Biological Effects Within Ranges of Chemical Concentrations in Marine and Estuarine Sediments," *Environmental Management.* Volume 19, Issue 1. p. 81-97, 1995.

Long, E.R. and L.G. Morgan., "The Potential Biological Effects of Sediment Sorbed Contaminants Tested in the National Status and Trends Program," NOAA Technical Memorandum NOS OMA 52, 1990.

Long, E.R. and R. Markel., "An Evaluation of the Extent and Magnitude of Biological Effects Associated with Chemical Contaminants in San Francisco Bay, California," NOAA Technical Memorandum NOS ORCA 64, 1992.

Naval Natural Resources Management Staff, "Natural Resource Management Plan," Prepared for the U.S. Department of the Navy, Western Division Naval Facilities Engineering Command in Cooperation with the U.S. Department of Agriculture Sediment Conservation Service. July, 1986.

Owen, C.J., R.P. Axler, D.R. Nordman, M. Schubauer-Berigan, K.B. Lodge, and J.P. Schubauer-Berigan, "Screening for PAHs by fluorescence Spectroscopy: a comparison of calibrations," Chemosphere, V.31:3345-3356, 1995.

Pacific Aerial Surveys, Aerial Photograph of NAS Alameda. Pacific Aerial Surveys, Oakland, CA. May 10, 1935.

PRC Environmental Management, Inc. (PRC), "Hydrogeology and Proposed Changes for Phase 5 of the RI/FS, Naval Air Station (NAS) Alameda, California," March, 1991.

PRC, "Naval Air Station Alameda Draft Ecological Assessment," 1994.

PRC, "Final Demonstration Plan For The Evaluation of Field Portable X-Ray Fluorescence Technologies, Superfund Innovative Technology Evaluation Program", Prepared for USEPA, Environmental Monitoring Systems Laboratory, Las Vegas, Nevada, 891193, March, 1995.

PRC, "Operable Unit 4 Follow-on Ecological Assessment Work Plan/Field Sampling Plan - Draft," Naval Air Station Alameda, Alameda, California. July, 1996a.

PRC, "Protocol for Contaminant Fate and Transport Modeling," Naval Air Station Alameda, Alameda, California. March 29, 1996b.

Radbruch, D.A., "Areal and Engineering Geology of the Oakland West Quadrangle, California," Miscellaneous Geologic Investigations MAP I-239. U.S. Geological Survey, 1957.

Rogers, J.D. and Figuers, S.H., "Engineering Geologic Site Characterization of the Greater Oakland-Alameda Area, Alameda and San Francisco Counties, California," Final Report to the National Science Foundation, December 30, 1991.

Russ, J.C., "Fundamentals of Energy Dispersive X-ray Analysis," Butterworths and Co., Ltd., p. 308, 1984.

Sanborn-Ferris Map Co. (Sanborn), Map of Alameda, California, 1897.

Skei, J.M., Price, N.B., Calvert, S.E. and Holtedahl, H., "The distribution of heavy metals in sediments of Sorfjord, West Norway," Wat. Air Sediment Pollut. 1, p.452-461, 1972.

Smith, L.H, "A Review of Circulation and Mixing Studies of San Francisco Bay, California," U.S. Geological Survey Circular 1015. Prepared in cooperation with the California State Water Resources Control Board, and the California Department of Water Resources, 1987.

Stallard, M.O., Apitz, S.E. and Dooley, C.A., "X-ray fluorescence spectrometry for field analysis of metals in marine sediments," Mar. Pollut. Bull. 31, 297-305, 1985.

Swift, R.P., "Evaluation of a field-portable X-ray fluorescence spectrometry method for use in remedial activities," Spectroscopy. 10, 31-35, 1995.

Tchan, Y.T., C.M. Chiou, “Bioassay of herbicides by bioluminescence. ACTA Phytopath. Acad. Sci. Hung (12): 3, 1977.

Turner Operating Manual, 1997

U.S. Department of the Navy (U.S. Navy), "Final Environmental Impact Statement for Proposed New Dredging," Western Division, Naval Facilities Engineering Command, Code 1833RH. August, 1990a.

U.S. Environmental Protection Agency (1993) “Data Quality Objectives Process for Superfund-Interim Final Guidance.” Office of Solid Waste and Emergency Response, Washington, D.C. EPA/540/R-93/071, 1993.

U.S. Environmental Protection Agency, "Long-Term Management Strategy for the Placement of Dredged Material in the San Francisco Bay Region - Draft," San Francisco Bay Conservation and Development Commission, San Francisco Bay Regional Water Quality Control Board, and California State Water Resources Control Board, April, 1996a.

U.S. Environmental Protection Agency, “Method 6200. Field Portable X-ray Fluorescence Spectrometry for the Determination of Elemental Concentrations in Soil and Sediment.” Office of Solid Waste and Emergency Response. Washington, D.C., 1996.

U.S. Navy, "Final Environmental Impact Statement - Candidate Base Closures/Realignment - San Francisco Bay Area," Western Division, Naval Facilities Engineering Command, San Bruno, CA, 1990b.

U.S. Navy, "Letter Regarding Sediment Transport in San Francisco Bay as it Relates to Sites Investigated Under the IRP at NAS Alameda," From Teresa Bernhard, Remedial Project Manager, Engineering Field Activity West, Naval Facilities Engineering Command. To NAS Alameda Regulatory Agency Personnel. June 13, 1996.

Table 1. Results of bioassays conducted using the QwikLite system. Bioluminescent dinoflagellates were used for both acute and chronic exposures.

TOXICANT	DURATION	IC50	Mysid LC50	TOXICANT
TBT	196 hrs	1.6 ppb	0.5 ppb	TBT
Silver	96 hrs	13 ppb	249 ppb	Silver
Copper Sulfate	96 hrs	23 ppb	120 - 140 ppb	Copper Sulfate
DBT	96 hrs	34 ppb	---	DBT
Lead	96 hrs	321 ppb	3130 ppb	Lead
Zinc	96 hrs	430 ppb	499 ppb	Zinc
Chromium	96 hrs	538 ppb	2030 ppb	Chromium
Cadmium	96 hrs	782 ppb	110 ppb	Cadmium

Table 2. Comparison of response to reference toxicant sodium dodecyl sulfate (SDS).

SPECIES	ENDPOINT	IC50/LC50 (mg/L)
Gonyaulax polyedra (Dinoflagellate-QwikLite)	Bioluminescence	1.4
Menidia beryllina (Silverside minnow)	Larval survival	1.8
Cyprinodon variegatus (Sheepshead minnow)	Larval survival	2.9
Arabacia punctulata (Sea Urchin)	Fertilization	3.2
Mysidopsis bahia (Mysid shrimp)	Survival	9.3

Table 3. Advantages and limitations of screening methods versus laboratory methods.

SCREENING	LAB ANALYSIS
Benefits	**Benefits**
‣ Rapid results can guide sampling locations	‣ Standard Methods that are very quantitative
‣ Potential for high data density for mapping	‣ Can often remove interferences
Limitations	**Limitations**
‣ Often non-specific	‣ Often blind sampling
‣ Semi-quantitative	‣ Long delays to results
‣ Matrix sensitive	‣ Expensive

Table 4. Pre-demonstration and demonstration results for Fe and Zn from FPXRF analyses (mg/kg, wet) and from certified analyses (mg/kg, dry), NAS Alameda, CA.

		FPXRF field			*Confirmatory*		*FPXRF field*			*Confirmatory*	
Date	Sample ID	Fe (mg/kg)	Q	stdev	Fe (mg/kg)	Q	Zn (mg/kg)	Q	stdev	Zn (mg/kg)	Q
07/97	SS01	19200			23000		231			383	
07/97	SS02	25800			49000		156			271	
07/97	SS05	30200			58200		105			161	
07/97	SS11	24500			30000		55			84.1	
07/97	SS14	30100			55500		120			166	
07/97	SS17	30000			57100		83			161	
07/97	SS19	29400			55400		65	b		156	
07/97	SS22	30700			54900		68	b		156	
07/97	SS24	27400			52300		164			210	
07/97	SS25	21300			35200		135			282	
07/97	SS26	23600			47000		205			383	
07/97	SS27	29200			51700		193			285	
07/97	SS28	25100			41200		182			250	
07/97	SS30	31000			59250		114			185	
07/97	SS31	28400			49100		149			284	
10/98	SL01	25900		701	37400		295		36	409	
10/98	SL02	27500		349	42300		201		19	256	
10/98	SL03	26800		213	43900		154		18	211	
10/98	SL04	26200		817	43200		134		13	189	
10/98	SL05	27100		137	43600		149		23	178	
10/98	SL06	12200		86	9040		180		19	175	
10/98	SL07	25400		304	40200		158		37	252	
10/98	SL08	26000		851	42300		139		35	178	
10/98	SL09	26800		395	38200		140		28	139	
10/98	SL10	26600		232	41100		115		4	175	
10/98	SL11	25900		40	42700		144		38	174	
	DL	**222**			**5**		**70**			**1**	

FPXRF field: on site measurement, mg/kg wet.

Confirmatory: standard laboratory analysis, mg/kg dry.

Stdev (standard deviation): standard deviation of triplicate measurements of sample

DL (detection limit): for FPXRF (six times the standard deviation of 15 replicate measurements of a quartz blank), for confirmatory analyses, DL is as reported by contractor.

Q (Qualifier): for FPXRF (a = value generated by instrument less than or equal to zero, b = value less than DL).

Table 5. Pre-demonstration and demonstration results for Cu and Pb from FPXRF analyses (mg/kg, wet) and from certified analyses (mg/kg, dry), NAS Alameda, CA.

		FPXRF field			***Confirmatory***		***FPXRF field***			***Confirmatory***	
Date	Sample ID	Cu (mg/kg)	Q	stdev	Cu (mg/kg)	Q	Pb (mg/kg)	Q	stdev	Pb (mg/kg)	Q
07/97	SS01	17	b		99.6		127			158	
07/97	SS02	18	b		158		77			146	
07/97	SS05	0	a		80.7		0			26.1	
07/97	SS11	0	a		31.1		0			19.2	
07/97	SS14	5	b		78.1		2			26.3	
07/97	SS17	0	a		77.8		0			25.6	
07/97	SS19	6	b		70.0		3			25.7	
07/97	SS22	0	a		70.0		5			24.8	
07/97	SS24	9	b		84.4		19			66.1	
07/97	SS25	17	b		125		78			186	
07/97	SS26	56	b		211		85			191	
07/97	SS27	95			294		25			60.3	
07/97	SS28	55	b		151		69			45.6	
07/97	SS30	7	b		103		1			37.1	
07/97	SS31	39	b		1120		34			78.7	
10/98	SL01	44	b	15	159		179		5	219	
10/98	SL02	28	b	25	114		68		1	155	
10/98	SL03	0	a	0	99		24	b	14	87.3	
10/98	SL04	0	a	0	83		20	b	2	62.8	
10/98	SL05	0	a	0	81		14	b	2	48	
10/98	SL06	0	a	0	71		181		91	202	
10/98	SL07	29	b	26	116		70		9	149	
10/98	SL08	0	a	0	81		23	b	8	61	
10/98	SL09	0	a	0	62		15	b	4	39.1	
10/98	SL10	0	a	0	79		8	b	8	61.7	
10/98	SL11	0	a	0	77		20	b	11	54.3	
	DL	**88**			**2**		**28**			**0.02**	

FPXRF field: on site measurement, mg/kg wet.

Confirmatory: standard laboratory analysis, mg/kg dry.

Stdev (standard deviation): standard deviation of triplicate measurements of sample

DL (detection limit): for FPXRF (six times the standard deviation of 15 replicate measurements of a quartz blank), for confirmatory analyses, DL is as reported by contractor.

Q (Qualifier): for FPXRF (a = value generated by instrument less than or equal to zero, b = value less than DL).

Table 6. Pre-demonstration and demonstration results for Cu (mg/kg), Pearl Harbor.

		FPXRF field			Confirmatory			FPXRF field			Confirmatory
Pre-Demo Date	Sample ID	Cu (mg/kg)	Q	Stdev	Cu (mg/kg)	Demo Date	Sample ID	Cu (mg/kg)	Q	Stdev	Cu (mg/kg)
1/20/98	1bx	54	b	NA	470	7/29/99	ML01	44	b		182
1/20/98	1cz	89		NA	242	7/29/99	ML02	43	b		107
1/20/98	1dz	323		NA	559	7/29/99	ML03	89			78
1/20/98	1ex	175		32	437	7/29/99	ML04	53	b		226
1/20/98	1gx	74	b	15	163	8/3/99	ML07	0	a		190
1/20/98	1gz	222		NA	309	8/3/99	ML08	50	b		125
1/20/98	1iz	212		24	568	8/3/99	ML09	89			98
1/20/98	1jy	93		NA	106	8/3/99	ML11	80	b	29	110
1/20/98	1kx	639		20	1889	8/3/99	ML12	59	b		190
1/20/98	1lx	368		8	689	8/3/99	ML14	37	b	36	182
1/20/98	1nz	81	b	NA	179	8/3/99	ML15	97			92
1/20/98	1pz	76	b	10	62	8/3/99	BP04	194			161
1/20/98	2ay	95		12	75	8/3/99	BP05	132			150
1/20/98	2az	9	b	NA	50	8/3/99	BP06	122		44	256
1/20/98	2bx	204		NA	295	8/3/99	BP07	214			199
1/20/98	2dz	72	b	NA	28	8/3/99	BP08	112			237
1/20/98	2hx	131		NA	400	8/3/99	BP09	139			184
1/20/98	2ix	154		NA	178	8/3/99	BP10	161		71	235
1/20/98	2iz	67	b	NA	120	8/3/99	BP11	166			187
1/20/98	2jx	53	b	NA	17	8/3/99	BP12	33	b		38
1/20/98	2mx	139		19	208	8/3/99	BP13	170		34	254
1/20/98	2nx	24	b	NA	8	8/3/99	BP14	75	b		81
1/20/98	3dx	7	b	NA	134	8/3/99	BP15	19	b		31
1/20/98	3ex	117		NA	163						
1/20/98	3hz	53	b	NA	60						
1/20/98	3iy	88		16	26						
1/20/98	3ly	19	b	NA	84						
1/20/98	4az	89		NA	303						
1/20/98	4bx	91		11	154						
1/20/98	4dz	31	b	NA	20						
1/20/98	4ex	97		NA	55						
1/20/98	4iy	164		NA	80						
1/20/98	5ax	122		NA	83						
1/20/98	5cx	52	b	NA	150						
1/20/98	5cy	160		NA	168						
1/20/98	5ex	0	a	NA	27						
1/20/98	5iz	119		NA	98						
1/20/98	5mx	56	b	NA	125						
1/20/98	5mz	198		86	198						
1/20/98	5nx	127		NA	102						
1/20/98	5rx	155		18	108						
1/20/98	5sy	82	b	NA	76						
DL		**88**						**88**			

Table 7. Pre-demonstration and demonstration results for Zn (mg/kg), Pearl Harbor.

		FPXRF lab			*Confirmatory*			*FPXRF lab*			*Confirmatory*
Date	Site Label	Zn (mg/kg)	Q	Stdev	Zn (mg/kg)	Date	Sample ID	Zn (mg/kg)	Q	stdev	Zn (mg/kg)
1/20/98	1bx	171		NA	445	7/29/99	ML01	164			387
1/20/98	1cz	176		NA	343	7/29/99	ML02	130			188
1/20/98	1dz	306		NA	546	7/29/99	ML03	205			197
1/20/98	1ex	249		32	425	7/29/99	ML04	152			401
1/20/98	1gx	150		25	233	8/3/99	ML07	184			335
1/20/98	1gz	203		NA	312	8/3/99	ML08	139			261
1/20/98	1iz	329		23	698	8/3/99	ML09	120			233
1/20/98	1jy	82		NA	149	8/3/99	ML11	113		68	237.5
1/20/98	1kx	370		24	725	8/3/99	ML12	198			340
1/20/98	1lx	428		46	693	8/3/99	ML14	148		8	308
1/20/98	1nz	132		NA	222	8/3/99	ML15	141			241
1/20/98	1pz	106		26	137	8/3/99	BP04	249			332
1/20/98	2ay	131		14	141	8/3/99	BP05	269			760
1/20/98	2az	118		NA	172	8/3/99	BP06	200		2	338
1/20/98	2bx	121		NA	235	8/3/99	BP07	202			345
1/20/98	2dz	67	b	NA	69	8/3/99	BP08	167			388
1/20/98	2hx	116		NA	281	8/3/99	BP09	211			1330
1/20/98	2ix	126		NA	192	8/3/99	BP10	263		26	487
1/20/98	2iz	459		NA	311	8/3/99	BP11	322			891
1/20/98	2jx	80		NA	65	8/3/99	BP12	81			88
1/20/98	2mx	239		21	262	8/3/99	BP13	387		28	1570
1/20/98	2nx	62	b	NA	37	8/3/99	BP14	177			209
1/20/98	3dx	74		NA	243	8/3/99	BP15	51	b		86
1/20/98	3ex	124		NA	177						
1/20/98	3hz	158		NA	163						
1/20/98	3iy	118		11	68						
1/20/98	3ly	89		NA	161						
1/20/98	4az	139		NA	393						
1/20/98	4bx	130		21	228						
1/20/98	4dz	116		NA	91						
1/20/98	4ex	91		NA	95						
1/20/98	4iy	183		NA	162						
1/20/98	5ax	137		NA	201						
1/20/98	5cx	130		NA	256						
1/20/98	5cy	235		NA	407						
1/20/98	5ex	74		NA	78						
1/20/98	5iz	157		NA	175						
1/20/98	5mx	199		NA	265						
1/20/98	5mz	316		25	368						
1/20/98	5nx	137		NA	227						
1/20/98	5rx	296		33	329						
1/20/98	5sy	156		NA	168						
DL		**70**						**70**			

Table 8. Pre-demonstration and demonstration results for Pb (mg/kg), Pearl Harbor.

		FPXRF lab			*Confirmatory*			*FPXRF lab*			*Confirmatory*
Date	Site Label	Pb (mg/kg)	Q	Stdev	Pb (mg/kg)	Date	Sample ID	Pb (mg/kg)	Q	stdev	Pb (mg/kg)
1/20/98	1bx	28		NA	200	7/29/99	ML01	4	b		32.9
1/20/98	1cz	96		NA	244	7/29/99	ML02	0	a		7.8
1/20/98	1dz	216		NA	429	7/29/99	ML03	0	a		9.4
1/20/98	1ex	117		11	306	7/29/99	ML04	0	a		115
1/20/98	1gx	40		15	97	8/3/99	ML07	0	a		33.3
1/20/98	1gz	47		NA	120	8/3/99	ML08	0	a		16.5
1/20/98	1iz	61		28	183	8/3/99	ML09	0	a		20.9
1/20/98	1jy	13	b	NA	51	8/3/99	ML11	0	a	0	11.15
1/20/98	1kx	138		16	459	8/3/99	ML12	0	a		41.1
1/20/98	1lx	97		6	239	8/3/99	ML14	3	b	3	29.0
1/20/98	1nz	0	a	NA	27	8/3/99	ML15	0	a		13.1
1/20/98	1pz	9	b	6	31	8/3/99	BP04	25	b		341
1/20/98	2ay	0	a	0	39	8/3/99	BP05	70			256
1/20/98	2az	17	b	NA	39	8/3/99	BP06	134		6	55.5
1/20/98	2bx	8	b	NA	62	8/3/99	BP07	19	b		99.3
1/20/98	2dz	0	a	NA	15	8/3/99	BP08	0	a		119
1/20/98	2hx	29		NA	116	8/3/99	BP09	21	b		236
1/20/98	2ix	28		NA	44	8/3/99	BP10	49		3	143
1/20/98	2iz	60		NA	85	8/3/99	BP11	93			173
1/20/98	2jx	15	b	NA	38	8/3/99	BP12	0	a		30.3
1/20/98	2mx	24	b	4	123	8/3/99	BP13	115		14	261
1/20/98	2nx	0	a	NA	3	8/3/99	BP14	31			76.1
1/20/98	3dx	4	b	NA	40	8/3/99	BP15	0	a		25.6
1/20/98	3ex	4	b	NA	22						
1/20/98	3hz	0	a	NA	22						
1/20/98	3iy	0	a	0	8						
1/20/98	3ly	0	a	NA	43						
1/20/98	4az	1	b	NA	45						
1/20/98	4bx	5	b	7	31						
1/20/98	4dz	12	b	NA	2						
1/20/98	4ex	4	b	NA	3						
1/20/98	4iy	0	a	NA	13						
1/20/98	5ax	0	a	NA	28						
1/20/98	5cx	9	b	NA	50						
1/20/98	5cy	10	b	NA	58						
1/20/98	5ex	0	a	NA	11						
1/20/98	5iz	0	a	NA	41						
1/20/98	5mx	41		NA	60						
1/20/98	5mz	76		12	134						
1/20/98	5nx	26	b	NA	53						
1/20/98	5rx	0	a	0	46						
1/20/98	5sy	14	b	NA	27						
DL		**28**						**28**			

Table 9. Pre-demonstration and demonstration results for Fe (mg/kg), Pearl Harbor.

		FPXRF lab			*Confirmatory*			*FPXRF lab*			*Confirmatory*
Date	Site Label	Fe (mg/kg)	Q	Stdev	Fe (mg/kg)	Date	Sample ID	Fe (mg/kg)	Q	stdev	Fe (mg/kg)
1/20/98	1bx	23800		NA	69978	7/29/99	ML01	34880			87200
1/20/98	1cz	27300		NA	31830	7/29/99	ML02	64441			101700
1/20/98	1dz	27300		NA	45653	7/29/99	ML03	67315			81400
1/20/98	1ex	23800		1168	41696	7/29/99	ML04	36604			88700
1/20/98	1gx	24600		493	38462	8/3/99	ML07	40167			94800
1/20/98	1gz	26000		NA	30748	8/3/99	ML08	65796			107100
1/20/98	1iz	21500		808	47204	8/3/99	ML09	67473			84600
1/20/98	1jy	27600		NA	39138	8/3/99	ML11	68760		846	109250
1/20/98	1kx	19900		529	34589	8/3/99	ML12	55885			91900
1/20/98	1lx	23400		208	41676	8/3/99	ML14	46802		638	99500
1/20/98	1nz	58500		NA	87803	8/3/99	ML15	58016			77300
1/20/98	1pz	23600		551	19755	8/3/99	BP04	12500			15300
1/20/98	2ay	22900		551	36501	8/3/99	BP05	11600			14800
1/20/98	2az	23100		NA	31098	8/3/99	BP06	11300		320	11700
1/20/98	2bx	28200		NA	44569	8/3/99	BP07	13700			21500
1/20/98	2dz	10500		NA	10457	8/3/99	BP08	12200			18100
1/20/98	2hx	14300		NA	26596	8/3/99	BP09	13000			18100
1/20/98	2ix	11300		NA	9389	8/3/99	BP10	15400		719	23300
1/20/98	2iz	12000		NA	14013	8/3/99	BP11	15300			24600
1/20/98	2jx	8900		NA	7101	8/3/99	BP12	9800			13500
1/20/98	2mx	16800		100	24463	8/3/99	BP13	15500		354	26100
1/20/98	2nx	6400		NA	6406	8/3/99	BP14	12500			19000
1/20/98	3dx	46000		NA	125353	8/3/99	BP15	8640			11600
1/20/98	3ex	80400		NA	121106						
1/20/98	3hz	72700		NA	94667						
1/20/98	3iy	37200		755	33333						
1/20/98	3ly	26300		NA	34493						
1/20/98	4az	39000		NA	91210						
1/20/98	4bx	54300		252	82283						
1/20/98	4dz	29500		NA	20457						
1/20/98	4ex	82400		NA	48650						
1/20/98	4iy	83500		NA	79906						
1/20/98	5ax	68000		NA	56035						
1/20/98	5cx	47800		NA	75441						
1/20/98	5cy	72300		NA	139000						
1/20/98	5ex	37400		NA	32241						
1/20/98	5iz	76300		NA	93600						
1/20/98	5mx	51800		NA	48872						
1/20/98	5mz	74500		6422	75957						
1/20/98	5nx	53700		NA	75700						
1/20/98	5rx	84700		265	128000						
1/20/98	5sy	41200		NA	48012						
DL		**222**						**222**			

Table 10. Results of linear regressions of Fe, Zn, Cu and Pb data, Demonstration #1 (NAS Alameda).

Data Set	Coefficient of Determination (R^2)	Slope (m)	Intercept (b)	Number of Samples (n)
Fe (0-60000 mg/kg)	0.86	0.3252	11949	26
Zn (0-450 mg/kg)	0.71	0.5552	24.61	26
Cu (0-1200 mg/kg)	0.18	0.0496	9.15	25
Cu (0-300 mg/kg)	0.87	0.4102	-28.14	26
Pb (0-250 mg/kg)	0.81	0.7302	-19.22	26

Table 11. On site and laboratory FPXRF results and certified laboratory results for Fe, Zn, Cu, and Pb from Demo #1 (NAS Alameda). (NA: not analyzed, ND: non-detect, DL: detection limit)

	in-lab	*on-site*	*certified*	*in-lab*	*on-site*	*certified*	*in-lab*	*on-site*	*certified*	*in-lab*	*on-site*	*certified*
Sample ID	Fe (ppm)	Fe (ppm)	Fe (ppm)	Zn (ppm)	Zn (ppm)	Zn (ppm)	Cu (ppm)	Cu (ppm)	Cu (ppm)	Pb (ppm)	Pb (ppm)	Pb (ppm
BWB01	25000	NA	40000	106	NA	145	ND	NA	66	11	NA	32.2
BWB02	25600	NA	39100	150	NA	141	ND	NA	64	ND	NA	32.5
BWB03	25900	NA	35300	139	NA	137	ND	NA	56	25	NA	36.7
BWB04	25200	NA	38100	80	NA	135	42	NA	58	0	NA	31.9
BWB05	25100	NA	39500	99	NA	144	ND	NA	60	0	NA	33.5
BWB06	26100	NA	33800	99	NA	125	ND	NA	55	0	NA	33.5
PA01	22900	NA	36400	178	NA	277	68	NA	184	66	NA	155
PA02	24000	NA	32600	124	NA	197	16	NA	161	29	NA	68.2
PA03	23400	NA	31700	153	NA	115	ND	NA	87	ND	NA	54.6
PA04	26700	NA	45000	88	NA	161	ND	NA	95	ND	NA	53.3
PA05	26300	NA	46400	72	NA	143	ND	NA	76	ND	NA	28.7
PA06	25100	NA	40700	144	NA	222	ND	NA	153	19	NA	82.4
RL01	23700	NA	29000	88	NA	80	ND	NA	28	ND	NA	15.5
RL02	27900	NA	36100	102	NA	110	ND	NA	45	ND	NA	24
RL03	23800	NA	19400	98	NA	55	ND	NA	21	ND	NA	9.71
RL04	26500	NA	29700	55	NA	94	ND	NA	41	ND	NA	20.6
RL05	27000	NA	32500	86	NA	103	ND	NA	43	6	NA	22.8
RL06	28000	NA	36700	81	NA	112	ND	NA	45	ND	NA	23.2
SL01	26800	25900	37400	236	295	409	26	44	159	171	179	219
SL02	27600	27500	42300	121	201	256	27	28	114	70	68	155
SL03	26700	26800	43900	142	154	211	55	ND	99	29	24	87.3
SL04	27800	26200	43200	136	134	189	55	ND	83	16	20	62.8
SL05	27000	27100	43600	120	149	178	27	ND	81	21	14	48
SL06	11700	12200	9040	203	180	175	101	ND	71	270	181	202
SL07	25700	25400	40200	204	158	252	37	29	116	71	70	149
SL08	25800	26000	42300	131	139	178	24	ND	81	20	23	61
SL09	27700	26800	38200	136	140	139	33	ND	62	15	15	39.1
SL10	26900	26600	41100	109	115	175	50	ND	79	20	8	61.7
SL11	25700	25900	42700	150	144	174	53	ND	77	20	20	54.3
D.L.	**222**	**222**		**70**	**70**		**88**	**88**		**28**	**28**	

Table 12. Results of linear regressions of Fe, Zn, Cu and Pb data, Demonstration #2 (Pearl Harbor).

Data Set	Coefficient of Determination (R^2)	Slope (m)	Intercept (b)	Number of Samples (n)
Fe (0-15 %)	0.70	0.5497	0.7445	66
Zn (0-2000 mg/kg)	0.46	0.2267	104.15	66
Zn (0-1000 mg/kg)	0.59	0.3760	65.769	64
Cu (0-2000 mg/kg)	0.78	0.3239	49.098	66
Cu (0-750 mg/kg)	0.46	0.3550	44.126	65
Pb (0-500 mg/kg)	0.64	0.3365	-2.338	66

Table 13. Field duplicate analyses with precision determined by relative percent difference (RPD).

		FPXRF	*Certified*	*FPXRF*	*Certified*	*FPXRF*	*Certified*	*FPXRF*	*Certified*
Demonstration	Sample ID	Fe	Fe	Zn	Zn	Cu	Cu	Pb	Pb
Demo #1	SL08	26000	42300	139	178	0	81	23	61
Demo #1	SL11	25900	42700	144	174	0	77	20	54.3
	RPD	0.10	0.24	0.83	0.57	NC	1.27	3.36	2.91
Demo #2	BP07	13652	NA	202	NA	214	NA	19	NA
Demo #2	BP18	13167	NA	330	NA	185	NA	30	NA
	RPD	0.90	NA	12	NA	4	NA	11	NA
Demo #2	BP11	15312	NA	322	NA	166	NA	93	NA
Demo #2	BP19	14530	NA	338	NA	187	NA	54	NA
	RPD	1.31	NA	1	NA	3	NA	13	NA

NA = not analyzed

Table 14. Instrument precision analyses with precision determined by percent relative standard deviation (%RSD)

		FPXRF	***replicate***	***FPXRF***	***replicate***	***FPXRF***	***replicate***	***FPXRF***	***replicate***
Demonstration	Sample ID	Fe (%)	%RSD	Zn (mg/kg)	%RSD	Pb (mg/kg)	%RSD	Cu (mg/kg)	%RSD
Demo #1	SS02	2.69	0.58	181	18	117	9	29	41
Demo #1	SS06	3.03	0.85	91	5	3	87	0	NC
Demo #1	SS19	2.94	0.22	65	30	3	43	6	153
Demo #1	SS21	2.92	1.16	123	6	2	58	21	28
Demo #1	SS25	2.13	1.25	135	15	78	2	17	127
Demo #1	SS26	2.36	0.72	205	4	85	8	56	13
Demo #1	SS27	2.92	1.48	193	3	25	9	95	8
Demo #1	SS31	2.84	0.43	149	5	34	19	39	56
Demo #1	SS33	1.90	0.20	270	3	204	2	26	50
Demo #1	BWB05	2.49	0.78	100	20	0	NC	0	NC
Demo #1	PA06	2.49	1.28	138	4	25	101	18	49
Demo #1	RL03	2.33	1.82	75	32	0	NC	0	NC
Demo #2	ML11	6.88	1.23	113	60	0	NC	80	36
Demo #2	BP6	1.13	2.83	200	1	134	4	122	36
Demo #2	BP13	1.55	2.28	387	7	115	12	170	20
Demo #2	BP22	1.43	2.07	192	8	32	60	133	20

NC: not calculated

Table 15. Sample heterogeneity precision analyses with precision determined by standard deviation (SD) and percent relative standard deviation (% RSD).

Demo	ID	% Fe	SD	% RSD	Zn ppm	SD	% RSD	Pb ppm	SD	% RSD	Cu ppm	SD	% RSD
#1	SL01	2.59	0.070	2.71	295	36	12	179	5	3	44	15	34
#1	SL02	2.75	0.035	1.27	201	19	9	68	1	2	28	25	88
#1	SL03	2.68	0.021	0.79	154	18	11	24	14	57	0	NC	NC
#1	SL04	2.62	0.082	3.11	134	13	10	20	2	10	0	NC	NC
#1	SL05	2.71	0.014	0.51	149	23	16	14	2	16	0	NC	NC
#1	SL06	1.22	0.009	0.71	180	19	10	181	91	50	0	NC	NC
#1	SL07	2.54	0.030	1.20	158	37	23	70	9	13	29	26	90
#1	SL08	2.60	0.085	3.28	139	35	25	23	8	35	0	NC	NC
#1	SL09	2.68	0.039	1.47	140	28	20	15	4	26	0	NC	NC
#1	SL10	2.66	0.023	0.87	115	4	4	8	8	93	0	NC	NC
#1	SL11	2.59	0.004	0.16	144	38	27	20	11	53	0	NC	NC
#2	ML05	3.61	0.028	0.78	182	15	8	0	NC	NC	60	13	21

NC = not calculated

Table 16. Performance evaluation samples analyzed during Demonstration #1 and Demonstration #2 as an indicator of instrument accuracy and stability.

SRM	Fe (mg/kg)	Zn (mg/kg)	Pb (mg/kg)	Cu (mg/kg)
PACS-1 (n=15)	36369	574	266	269
SD	987	24	13	46
Certified	49000	824	404	452
% Yield	74	70	66	59
N2704 (n=2)	35500	378	93	23
SD	0	19	23	2
Certified	41100	438	161	98.6
% Yield	86	89	48	25
N2710 (n=2)	28200	5450	4340	2330
SD	185	10	4	43
Certified	33800	6952	5532	2950
% Yield	83	78	78	79
N2711 (n=3)	22800	336	918	43
SD	440	63	21	53
Certified	28900	350	1162	114
% Yield	79	96	79	38

Table 17. Predemonstration and Demonstration Samples From Pier Area in Demonstration #1.

Date	Site Label	UVF [Intensity]	Stdev	%RSD	Confirmatory PAH (μg/kg)
04-Nov-98	PA01	1289			64016
04-Nov-98	PA02	887			41785
03-Nov-98	PA03	448			14694
03-Nov-98	PA04	260			5559
03-Nov-98	PA05	76			4138
04-Nov-98	PA06	1083	206	19.02%	93210
08-Dec-97	SSP11	1680			122282
08-Dec-97	SSP12	848			106979
08-Dec-97	SSP13	672	36	5.36%	41246
08-Dec-97	SSP14	1194			78662
08-Dec-97	SSP15	415	40	9.65%	38385
05-Nov-98	BWB1	170			6700
05-Nov-98	BWB2	125			2122
04-Nov-98	BWB3	126			1613
03-Nov-98	BWB4	97			1335
03-Nov-98	BWB5	100			1236
04-Nov-98	BWB6	140	13	9.29%	1683
08-Dec-97	BWB16	137	14	10.25%	11584
08-Dec-97	BWB17	100	11	10.95%	3450
08-Dec-97	BWB18	53			1463
08-Dec-97	BWB19	86			527
08-Dec-97	BWB20	74			173

Table 18. Predemonstration and Demonstration Samples From Seaplane Lagoon in Demonstration #1.

Date	*Site Label*	*UVF [Intensity]*	*Stdev*	*% RSD*	*Confirmatory PAH (μg/kg)*
27-Oct-98	SL01	1300			10871
27-Oct-98	SL02	485			3672
27-Oct-98	SL03	212			2992
27-Oct-98	SL04	110			2633
27-Oct-98	SL05	76			2278
28-Oct-98	SL06	690			10452
28-Oct-98	SL07	313			3137
28-Oct-98	SL08	91			2360
28-Oct-98	SL09	97			2367
28-Oct-98	SL10	107			2180
28-Oct-98	SL11	102	11	10.78%	1896
08-Dec-97	SPL01	831	60	7.22%	16973
08-Dec-97	SPL03	114			1750
08-Dec-97	SPL04	88			1154
08-Dec-97	SPL05	94			1664
08-Dec-97	SPL06	592	13	2.20%	4828
08-Dec-97	SPL07	697	50	7.17%	15012
08-Dec-97	SPL08	91			923
08-Dec-97	SPL09	75			1536
08-Dec-97	SPL10	77			1679
08-Dec-97	SPL02	836	25	2.99%	8695

Table 19. Predemonstration and Demonstration Samples from Demonstration #2.

Date	Site Label	UVF [Intensity]	Stdev	% RSD	Confirmatory PAH (μg/kg)
08-Feb-99	BP04	719			81241
08-Feb-99	BP05	1214			100254
08-Feb-99	BP06	398			32168
08-Feb-99	BP07	1111	91	8.19%	148603
08-Feb-99	BP08	838			82948
08-Feb-99	BP09	841			50635
08-Feb-99	BP10	991			80884
08-Feb-99	BP11	656			48510
08-Feb-99	BP12	245			7157
08-Feb-99	BP13	369			64680
08-Feb-99	BP14	473			19309
08-Feb-99	BP15	331			7064
09-Feb-99	BP17	1993			198867
09-Feb-99	BP18	740			82362
09-Feb-99	BP19	771	62	8.04%	45624
09-Feb-99	BP20	786			69598
09-Feb-99	BP21	731			47649
09-Feb-99	BP22	765			33892
09-Feb-99	BP23	1792			208740
02-Feb-99	ML07	119			1307
02-Feb-99	ML08	159			1062
02-Feb-99	ML09	106	11	10.39%	1630
02-Feb-99	ML12	317			2675

Table 20. Comparison of pre-demonstration Amphipod (Eohaustorius) toxicity data with QwikSed demonstration toxicity data at NAS Alameda for 25 samples.

Sample ID	QwikSed % Control at 25%	% Development-Amphipod	% Development Sea Urchin
BWB01	100	75	89
BWB02	100	61	87.8
BWB03	100	60	87.2
BWB04	70.51	47	91.8
BWB05	93.09	66	92.4
BWB06	100	n/a	n/a
PA01	64.05	15	79.4
PA02	99.3	19	63.2
PA03	64.78	32	88
PA04	100	62	84
PA05	100	53	88.8
PA06	100	n/a	n/a
SL1	95.64	73.8	89.8
SL2	84.54	50	95.2
SL3	100	60	87
SL4	91.44	59	77.4
SL5	79.91	69	77.2
SL6	92.21	90	81.4
SL7	100	58	93.8
SL8	100	64	92.2
SL9	100	70	95
SL10	100	67	91.8
SL11	n/a	n/a	n/a
RL01	71.45	8	89.5
RL02	78.61	41	85.6
RL03	100	83	91.2
RL04	100	50	93.8
RL05	100	58.8	96.4
RL06	100	n/a	n/a

Table 21. Comparison of pre-demonstration Amphipod (Eohaustorius) toxicity data with QwikSed demonstration toxicity data at NAS Alameda for 25 samples.

AMPHIPOD	**QwikSed Results**		
	TOXIC	**NOT TOXIC**	**Total**
TOXIC	5	1	6
NOT TOXIC	1	18	19
Total	6	19	25
Both toxic	5 / 25	20%	
Both Not toxic	18 / 25	72%	
Total	23 / 25	**92% agreement**	

Table 22. Comparison of the Sea Urchin Development toxicity data with QwikSed toxicity data at NAS Alameda for 25 samples.

SEAURCHIN	**QwikSed Results**		
	TOXIC	**NOT TOXIC**	**Total**
TOXIC	2	2	4
NOT TOXIC	4	17	21
Total	6	19	25
Both toxic	2 / 25	8%	
Both Not toxic	17 / 25	68%	
Total	19 / 25	**76% agreement**	

Table 23. Comparison of pre-demonstration Amphipod toxicity data with Sea Urchin development toxicity demonstration data at NAS Alameda for 25 samples.

SEAURCHIN	**QwikSed Results**		
	TOXIC	**NOT TOXIC**	**Total**
TOXIC	2	2	4
NOT TOXIC	5	16	21
Total	7	18	25
Both toxic	2 / 25	8%	
Both Not toxic	16 / 25	64%	
Total	18 / 25	**72% agreement**	

Table 24. Echinoderm Development (MEC Analytical) vs. QwikSed – Pearl Harbor, Hawaii.

*Note: QwikSed tests using *Ceratocorys horrida*

Sample ID	Echinoderm % Control at 100%	QwikSed % Control- at 25%
ML7	0	93.54
ML8	96.4	94.27
ML9	96.2	94.49
ML13	96.1	99
BP4	8.3	55
BP5	0	81.68
BP6	39.4	65.69
BP7	0	83.63
BP8	61.2	91.49
BP9	0	86.46
BP10	49.2	89.25
BP11	0.6	26.12
BP12	47.5	70.05
BP13	0	82.64
BP14	4.7	73.98
BP15	0	73.92
BP17	94.6	69.25
BP21	1.8	87.19

Table 25. Comparison of the Sea Urchin Development toxicity data with QwikSed toxicity data at Pearl Harbor.

SEAURCHIN	**QwikSed Results**		
	TOXIC	**NOT TOXIC**	**Total**
TOXIC	13	2	15
NOT TOXIC	1	2	3
Total	14	4	18
Both toxic	13 / 18	72%	
Both Not toxic	2 / 18	11%	
Total	15 / 18	**83% agreement**	

Table 26. Advantages and Disadvantages of XRF and GFAA (USEPA, 2000).

Field Portable XRF Advantages	GFAA Advantages
❑ Field Portable (< 30 lbs) ❑ Operated by battery ❑ Rapid Analysis (*e.g*. five min). ❑ Multi-element analysis with wide dynamic range (ppm – percent levels) ❑ Non-destructive analysis. ❑ No waste is generated ❑ Operators usually can be trained in one or two days. The software is menu-driven. ❑ Little or no sample preparation is required.	❑ Greater sensitivity and detection limits than other methods ❑ Direct analysis of some types of liquid samples ❑ Low spectral interference ❑ Very small sample size
Field Portable XRF Disadvantages	**GFAA Disadvantages**
❑ Detection limits for chromium are 200 mg/kg or higher. Action levels for some elements, such as arsenic or cadmium, ❑ may be lower than the detection limits of XRF. ❑ Concentrations of elements in different types of soil or matrices might change, causing interferences (*e.g*. As:Pb). Site-specific calibration standards can compensate for some of those effects. ❑ A specific license is required to operate some FPXRF instruments. The total cost of attending a radiation safety ❑ course, obtaining the necessary paperwork, and paying the fee for the license can range from $500 to $1,000. ❑ The Cd-109 source should be replaced every two years ($4,000 to $5,000). ❑ Instrument with Si(Li) detector will require liquid nitrogen and a dewar (aluminum container) to hold the liquid nitrogen. This requirement adds the time and cost of obtaining and handling liquid nitrogen to cool an instrument with a Si(Li) detector before analysis can be performed.	❑ Continued calibration of the instrument is a component of the overall quality control plan and should be performed by analyzing one mid-concentration standard after every 10 analyses. ❑ Limited dynamic range ❑ High matrix interference ❑ Sample preparation for metals in soil or sediments is somewhat extensive and may require the use of a mobile laboratory. ❑ 220-volt power source required

Table 27. Comparison of toxicity tests.

	Toxicity Test Type		
	Sea Urchin	**Amphipod**	**QwikSed**
Test Duration	3-4 days	14-15 days	1 day
Organism Source	-- Shipped from supplier --		In-house
Endpoint	-- Laval Development Survival --		Biolumin.
Test facility	-- Send to Lab --		On-site
Data Results	14-30 days	21 days	1 day
Cost / Sample	$ 1,000	$ 1,400	$ 200
Agreement Among Tests	$r^2 = 0.943$		
	76%, 83% in agreement		
		72% in agreement	**92% in agreement**

Table 28. Relative FPXRF analytical costs.

I. On Site FPXRF Costs (Continental US Example; *e.g.* Demo Site #1)

# Samples	*Purchase (Spectrace 9000)*	*Lease (Spectrace 9000)*	*Cost Per Sample (Certified)*
n = 30 (includes QAQC)	$8000 (instrument)	$3600 (two weeks +S/H)	
Supplies	$0 (supplies)	$0 (supplies)	
analysis time = 1 day	$775 (labor@$86/hr for 9 hr)	$775 (labor@$86/hr for 9 hr)	
mob/demob = 1 day	$775 (labor@$86/hr for 9 hr)	$775 (labor@$86/hr for 9 hr)	
per diem = 2.5 days	$188 (perdiem@75/day)	$188 (perdiem@75/day)	
travel (air)	$250 (airfare, gov't rate)	$250 (airfare, gov't rate)	
rental car = 2.5 days	$75 (gov't rate)	$75 (gov't rate)	
Validation Samples (20%)	$900 (at $150/sample for n=6)	$900 (at $150/sample for n=6)	
Cost	$2,034 per sample	$220 per sample	
Exclude Validation Cost	$2,003 per sample	$190 per sample	
Exclude Instrument Cost	$100 per sample	$100 per sample	
Exclude Instr. & Val. Cost	$70 per sample	$70 per sample	$150 - 300 per sample
n = 150 (includes QAQC)	$8000 (instrument)	$3600 (instrument)	
	$250 (supplies)	$250 (supplies)	
analysis time = 5 days	$3870 (labor@$86/hr for 40 hr)	$3870 (labor@$86/hr for 40 hr)	
mob/demob = 1 day	$775 (labor@$86/hr for 9 hr)	$775 (labor@$86/hr for 9 hr)	
per diem = 6.5 days	$488 (perdiem@75/day)	$488 (perdiem@75/day)	
travel (air)	$250 (airfare, gov't rate)	$250 (airfare, gov't rate)	
rental car	$195 (gov't rate)	$195 (gov't rate)	
Validation Samples (20%)	$4500 (at $150/sample for n=30)	$4500 (at $150/sample for n=30)	
Cost	$455 per sample	$93 per sample	
Exclude Validation Cost	$426 per sample	$63 per sample	
Exclude Instrument Cost	$69 per sample	$69 per sample	
Exclude Instr. & Val. Cost	$39 per sample	$39 per sample	$150 - 300 per sample

II. In House

# Samples	*Purchase[a] (Spectrace 9000)*	*Lease (Spectrace 9000)*	*Cost Per Sample (Certified)*
n = 30 (includes QAQC)	$8000 (instrument)	$3600 (two weeks +S/H)	
supplies[b]	$0 (supplies)	$0 (supplies)	
analysis time = 1 day	$775 (labor@$86/hr for 9 hr)	$775 (labor@$86/hr for 9 hr)	
Validation Samples (20%)	$900 (at $150/sample for n=6)	$900 (at $150/sample for n=6)	
Cost	$1,991 per sample	$178 per sample	
Exclude Validation Cost	$1,961 per sample	$148 per sample	
Exclude Instrument Cost	$58 per sample	$58 per sample	
Exclude Instr. & Val. Cost	$28 per sample	$28 per sample	$150 - 300 per sample
n = 150 (includes QAQC)	$8000 (instrument)	$3600 (instrument)	
	$250 (supplies)	$250 (supplies)	
analysis time = 5 days	$3870 (labor@$86/hr for 40 hr)	$3870 (labor@$86/hr for 40 hr)	
Validation Samples (20%)	$4500 (at $150/sample for n=30)	$4500 (at $150/sample for n=30)	
Cost	$444 per sample	$81 per sample	
Exclude Validation Cost	$414 per sample	$51 per sample	
Exclude Instrument Cost	$58 per sample	$58 per sample	
Exclude Instr. & Val. Cost	$28 per sample	$28 per sample	$150 - 300 per sample

Note: a: This purchase scenario is based on a unit purchase for a single project. The cost per sample for this scenario would decrease significantly of the purchased unit were used on multiple projects.
b: FPXRF Supplies: XRF sample cups, Mylar film, gloves, mixing rods, etc.
c: Sample shipment costs and data analysis/reporting costs are not included here

Table 29. Relative UVF analytical costs.

I. On Site UVF Costs (Continental US Example; *e.g.* Demo Site #1)

# Samples	*Purchase[a] (Turner Fluorometer)*	*Lease (Turner Fluorometer)*	*Cost Per Sample (Certified)*
n = 30 (includes QAQC)	9500 (instrument)	1200 (two weeks +S/H)	
supplies[b]	50 (supplies)	50 (supplies)	
analysis time = 1 day	775 (labor@$86/hr for 9 hr)	775 (labor@$86/hr for 9 hr)	
mob/demob = 1 day	775 (labor@$86/hr for 9 hr)	775 (labor@$86/hr for 9 hr)	
per diem = 2.5 days	188 (perdiem@75/day)	188 (perdiem@75/day)	
travel (air)	250 (airfare, gov't rate)	250 (airfare, gov't rate)	
rental car = 2.5 days	75 (gov't rate)	75 (gov't rate)	
Validation Samples (20%)	3000 (at $500/sample for n=6)	3000 (at $500/sample for n=6)	
Cost	$487 per sample	$210 per sample	
Exclude Validation Cost	$387 per sample	$110 per sample	
Exclude Instrument Cost	$170 per sample	$170 per sample	
Exclude Instr. & Val. Cost	$70 per sample	$70 per sample	$500 per sample
n = 150 (includes QAQC)	9500 (instrument)	1200 (instrument)	
	250 (supplies)	250 (supplies)	
analysis time = 5 days	3870 (labor@$86/hr for 40 hr)	3870 (labor@$86/hr for 40 hr)	
mob/demob = 1 day	775 (labor@$86/hr for 9 hr)	775 (labor@$86/hr for 9 hr)	
per diem = 6.5 days	488 (perdiem@75/day)	488 (perdiem@75/day)	
travel (air)	250 (airfare, gov't rate)	250 (airfare, gov't rate)	
rental car	195 (gov't rate)	195 (gov't rate)	
Validation Samples (20%)	15000 (at $500/sample for n=30)	15000 (at $500/sample for n=30)	
Cost	$202 per sample	$147 per sample	
Exclude Validation Cost	$102 per sample	$47 per sample	
Exclude Instrument Cost	$139 per sample	$139 per sample	
Exclude Instr. & Val. Cost	$39 per sample	$39 per sample	$500 per sample

II. In House

# Samples	*Purchase[a] (Turner Fluorometer)*	*Lease (Turner Fluorometer)*	*Cost Per Sample (Certified)*
n = 30 (includes QAQC)	9500 (instrument)	1200 (two weeks +S/H)	
supplies[b]	50 (supplies)	50 (supplies)	
analysis time = 1 day	775 (labor@$86/hr for 9 hr)	775 (labor@$86/hr for 9 hr)	
Validation Samples (20%)	3000 (at $500/sample for n=6)	3000 (at $500/sample for n=6)	
Cost	$444 per sample	$168 per sample	
Exclude Validation Cost	$344 per sample	$68 per sample	
Exclude Instrument Cost	$128 per sample	$128 per sample	
Exclude Instr. & Val. Cost	$28 per sample	$28 per sample	$500 per sample
n = 150 (includes QAQC)	9500 (instrument)	1200 (two weeks +S/H)	
	250 (supplies)	250 (supplies)	
analysis time = 5 days	3870 (labor@$86/hr for 40 hr)	3870 (labor@$86/hr for 40 hr)	
Validation Samples (20%)	15000 (at $500/sample for n=30)	15000 (at $500/sample for n=30)	
Cost	$191 per sample	$135 per sample	
Exclude Validation Cost	$91 per sample	$35 per sample	
Exclude Instrument Cost	$127 per sample	$127 per sample	
Exclude Instr. & Val. Cost	$27 per sample	$27 per sample	$500 per sample

Note:
a: This purchase scenario is based on a unit purchase for a single project. The cost per sample for this scenario would decrease significantly if the purchased unit were costed over multiple projects.
b: UVF Supplies: Hexane solvent, glassware, gloves, mixing rods, etc.
c: Sample shipment costs and data analysis/reporting costs are not included here.

Table 30. Relative QwikSed analytical costs.

I. On Site QwikSed Costs(Continental US Example; *e.g.* Demo Site #1) [a]

# Samples	*Purchase[a] (QwikSed Toxicity)*	*Lease (QwikSed)*	*Cost Per Sample (Certified)*
n = 30 (includes QAQC)	15,000 (instrument)	500 (two weeks +S/H)	
supplies[b]	50 (supplies)	50 (supplies)	
analysis time = 1 day	775 (labor@$86/hr for 9 hr)	775 (labor@$86/hr for 9 hr)	
mob/demob = 1 day	775 (labor@$86/hr for 9 hr)	775 (labor@$86/hr for 9 hr)	
per diem = 5 days	375 (perdiem@75/day)	375 (perdiem@75/day)	
travel (air)	250 (airfare, gov't rate)	250 (airfare, gov't rate)	
rental car = 5 days	150 (gov't rate)	150 (gov't rate)	
Validation Samples (20%)	6,000 (at $1000/sample for n=6)	6,000 (at $1000/sample for n=6)	
Cost	$778 per sample	$294 per sample	
Exclude Validation Cost	$578 per sample	$94 per sample	
Exclude Instrument Cost	$278 per sample	$278 per sample	
Exclude Instr. & Val. Cost	$78 per sample	$78 per sample	$1000 per sample
n = 150 (includes QAQC)	15,000 (instrument)	1000 (instrument)	
	250 (supplies)	250 (supplies)	
analysis time = 25 days	19,375 (labor@$86/hr for 40 hr)	19,375 (labor@$86/hr for 40 hr)	
mob/demob = 1 day	775 (labor@$86/hr for 9 hr)	775 (labor@$86/hr for 9 hr)	
per diem = 26.5 days	1988 (perdiem@75/day)	1988 (perdiem@75/day)	
travel (air)	250 (airfare, gov't rate)	250 (airfare, gov't rate)	
rental car	810 (gov't rate)	810 (gov't rate)	
Validation Samples (20%)	30,000 (at $1000/sample for n=30)30,000 (at $1000/sample for n=30	15000 (at $500/sample for n=30)	
Cost	$455 per sample	$363 per sample	
Exclude Validation Cost	$255 per sample	$163 per sample	
Exclude Instrument Cost	$355 per sample	$356 per sample	
Exclude Instr. & Val. Cost	$155 per sample	$156 per sample	$1000 per sample

II. In House

# Samples	*Purchase[a] (QwikSed Toxicity)*	*Lease (QwikSed)*	*Cost Per Sample (Certified)*
n = 30 (includes QAQC)	15,000 (instrument)	500 (two weeks +S/H)	
supplies[b]	50 (supplies)	50 (supplies)	
analysis time = 5 day	3875 (labor@$86/hr for 9 hr)	3875 (labor@$86/hr for 9 hr)	
Validation Samples (20%)	6000 (at $1000/sample for n=6)	6000 (at $1000/sample for n=6)	
Cost	$829 per sample	$346 per sample	
Exclude Validation Cost	$629 per sample	$145 per sample	
Exclude Instrument Cost	$329 per sample	$329 per sample	
Exclude Instr. & Val. Cost	$129 per sample	$129 per sample	$1000 per sample
n = 150 (includes QAQC)	15,000 (instrument)	1000 (Four weeks +S/H)	
	250 (supplies)	250 (supplies)	
analysis time = 25 days	19,375 (labor@$86/hr for 40 hr)	19,375 (labor@$86/hr for 40 hr)	
Validation Samples (20%)	30,000 (at $1000/sample for n=30)30,000 (at $1000/sample for n=30	15000 (at $500/sample for n=30)	
Cost	$431 per sample	$338 per sample	
Exclude Validation Cost	$230 per sample	$138 per sample	
Exclude Instrument Cost	$331 per sample	$331 per sample	
Exclude Instr. & Val. Cost	$131 per sample	$131 per sample	$1000 per sample

Note:
a: This purchase scenario is based on a unit purchase for a single project. The cost per sample for this scenario would decrease significantly if the purchased unit were costed over multiple projects.
b: QwikSed Supplies
c. Sample shipment costs and data analysis/reporting costs are not included here

Table 31. Predicted FPXRF, UVF and QwikSed annual operation and maintenance costs.

	Operation / Maintenance Costs			
	FPXRF	**UVF**	**QwikSed**	**Combined/General**
Consumables				--
Miscellaneous Laboratory Supplies (*e.g.,* gloves, stirring rods, cuvettes)	$150/yr	$250/yr	$150/yr	
Labor	$200/day[a]	$200/day[b]	$200/day[c]	--
Data Interpretation	--	--	--	$500- 1000/day

* -- Signifies Not Applicable.
[a] Estimated average progress is 40 samples/day.
[b] Estimated average progress is 20 samples/day.
[c] Estimated average progress is 10 samples/day.

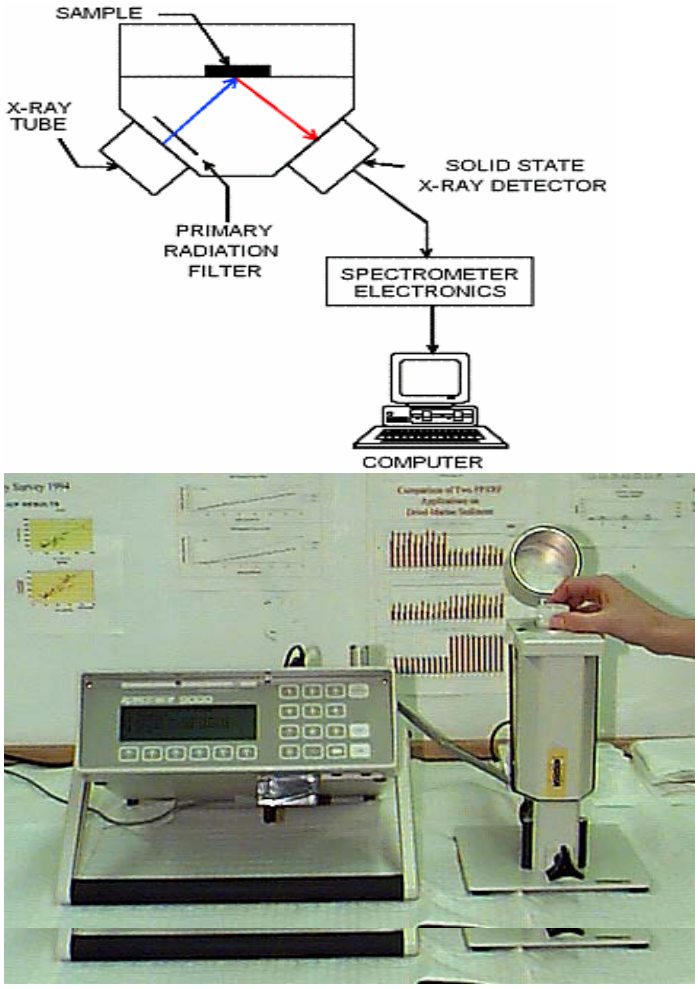


Figure 1. Generic schematic of an XRF system and photograph of a Spectrace 9000FPXRF unit (Spectrace Instruments).

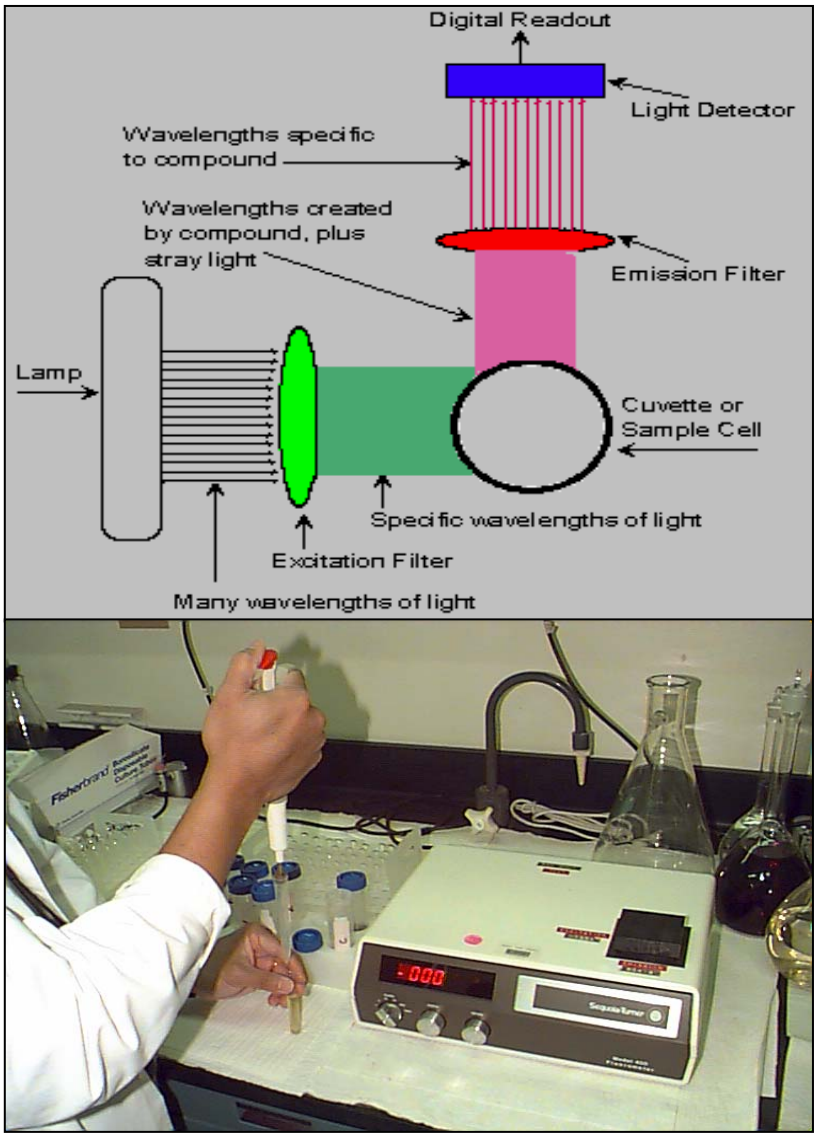


Figure 2. Generic schematic of the UVF system (top) and Photograph of the Turner UVF System (Turner Instruments).

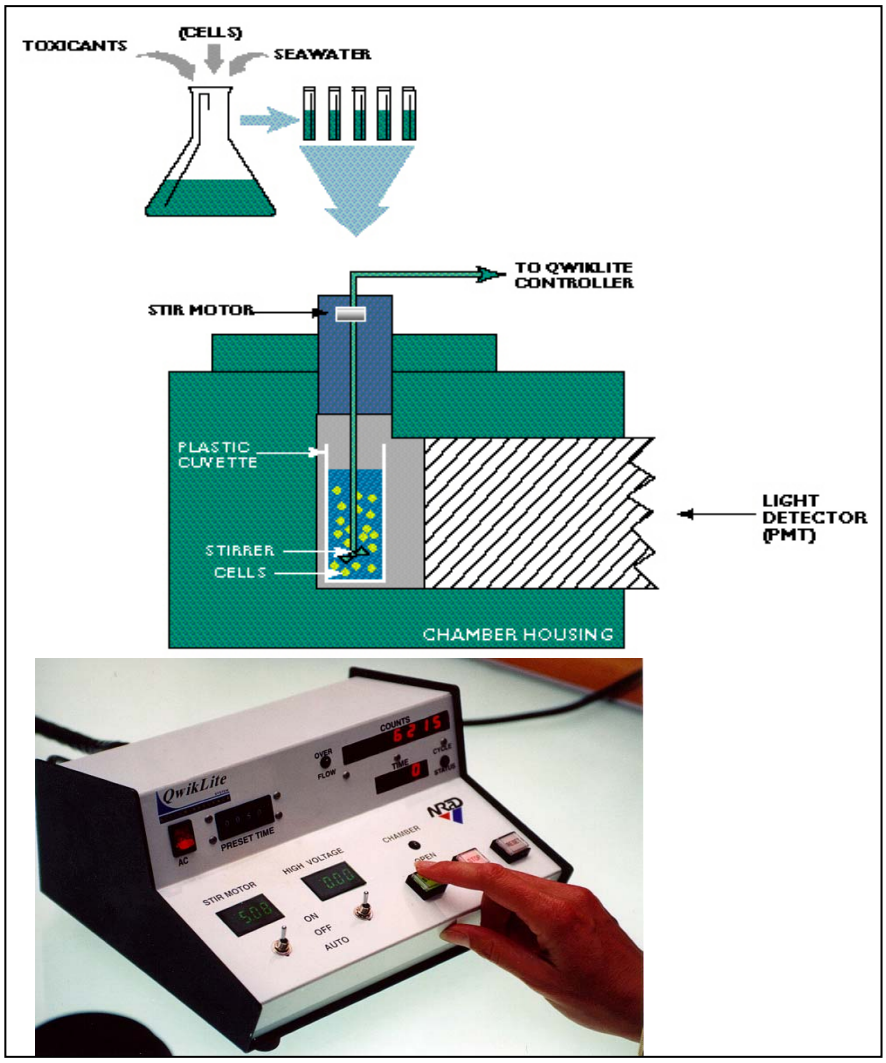


Figure 3. Schematic and photograph of the QwikSed System.

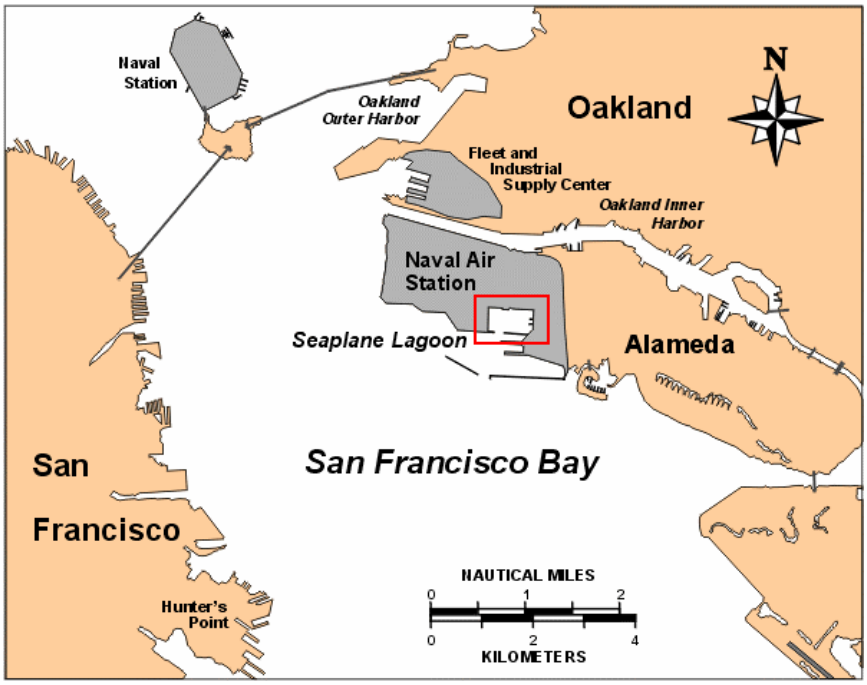


Figure 4. Map of Alameda Harbor including Seaplane Lagoon

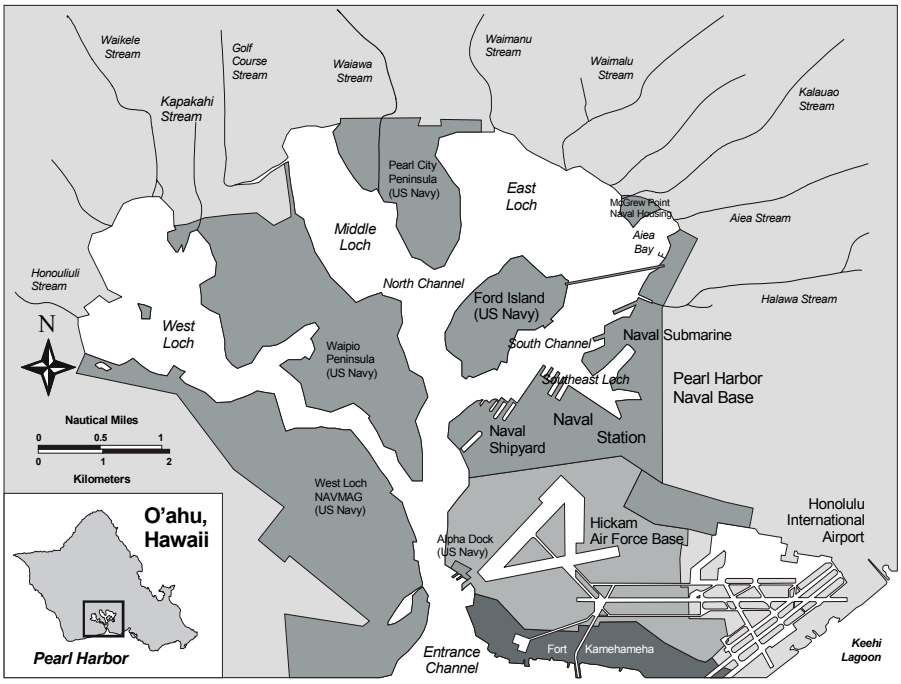


Figure 5. Pearl Harbor Naval Center.

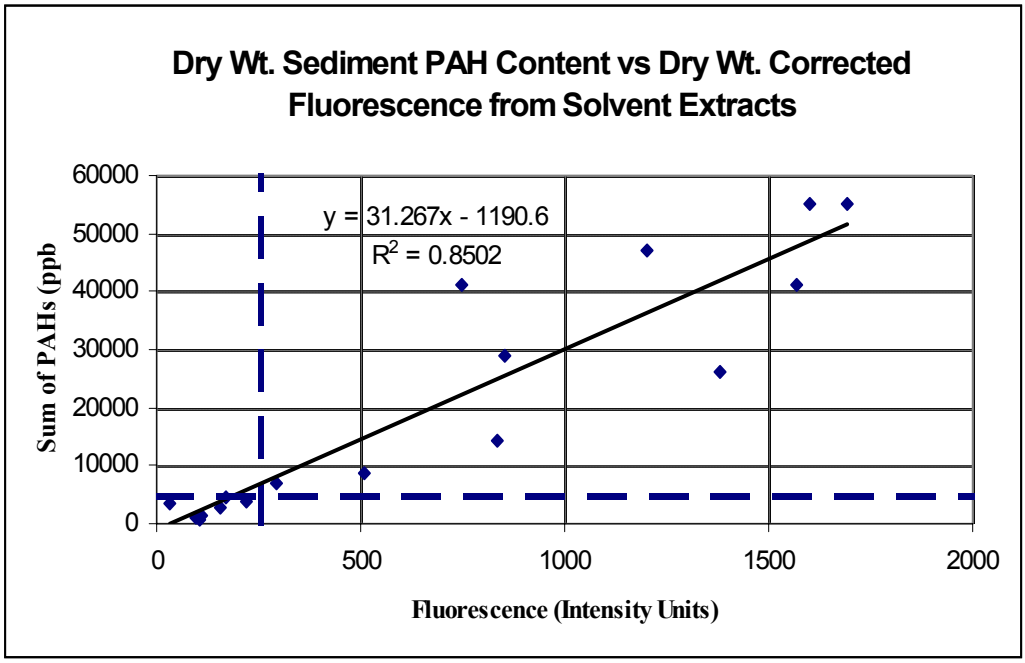
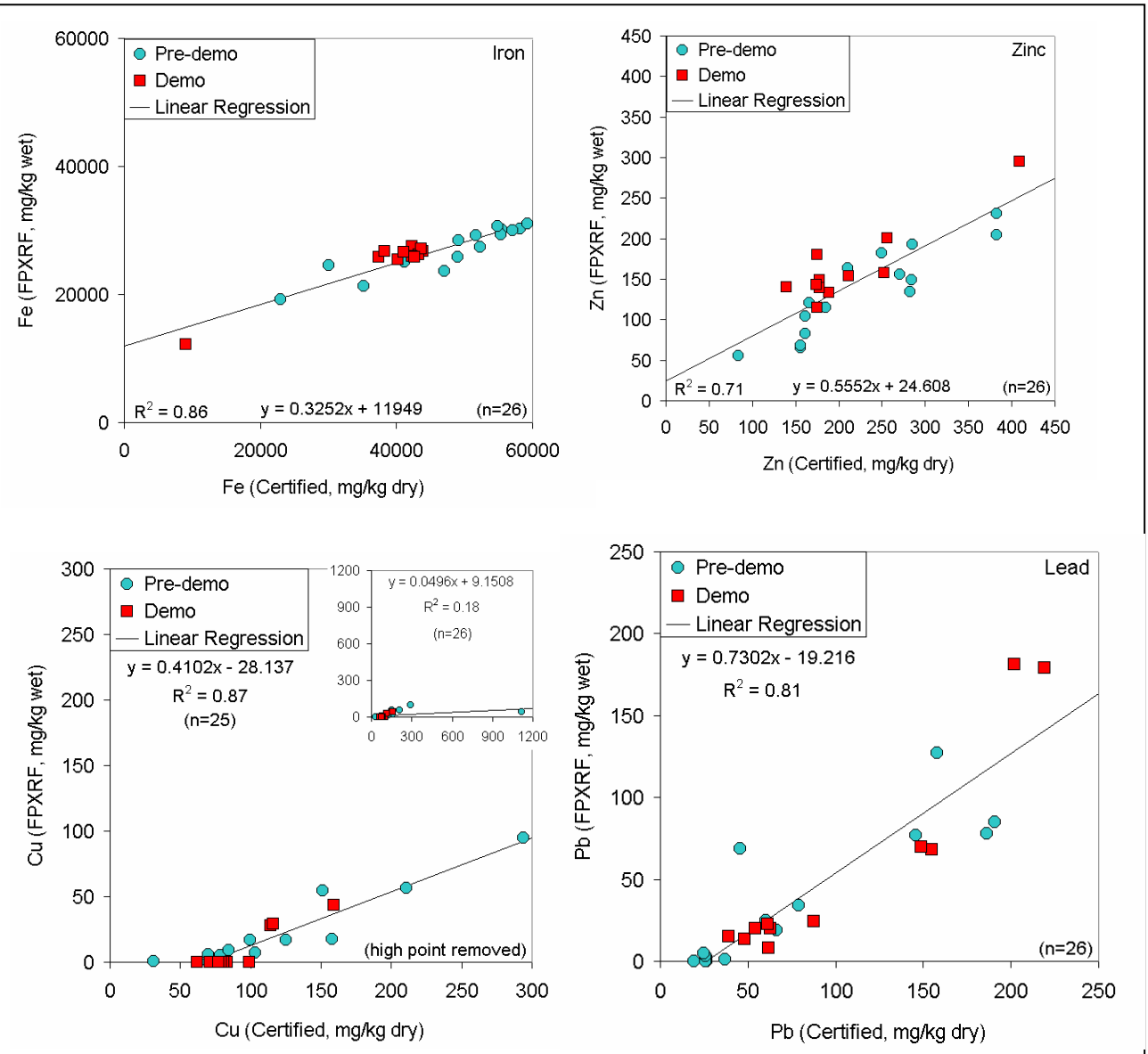


Figure 6. UVF pre-demonstration data divided into four Quadrants: Detect (D.), Under Detection Levels (U.), False Negative (F.N.), and False Positive (F.P.).

Figure 7. Spectrace 9000 FPXRF (probe, multi-channel analyzer, laptop computer).

Figure 8. On site analysis by FPXRF, Demo #2.



Inset of graph in Cu panel shows all samples, and blowup shows all samples in lower concentration range.

Circles represent samples collected from pre-demonstration at NAS Alameda – Southshore Pier area and squares represent samples collected during demonstration at NAS Alameda – Seaplane Lagoon.

Figure 9. FPXRF FE, Zn, Cu, and Pb results plotted against results from standard methods.

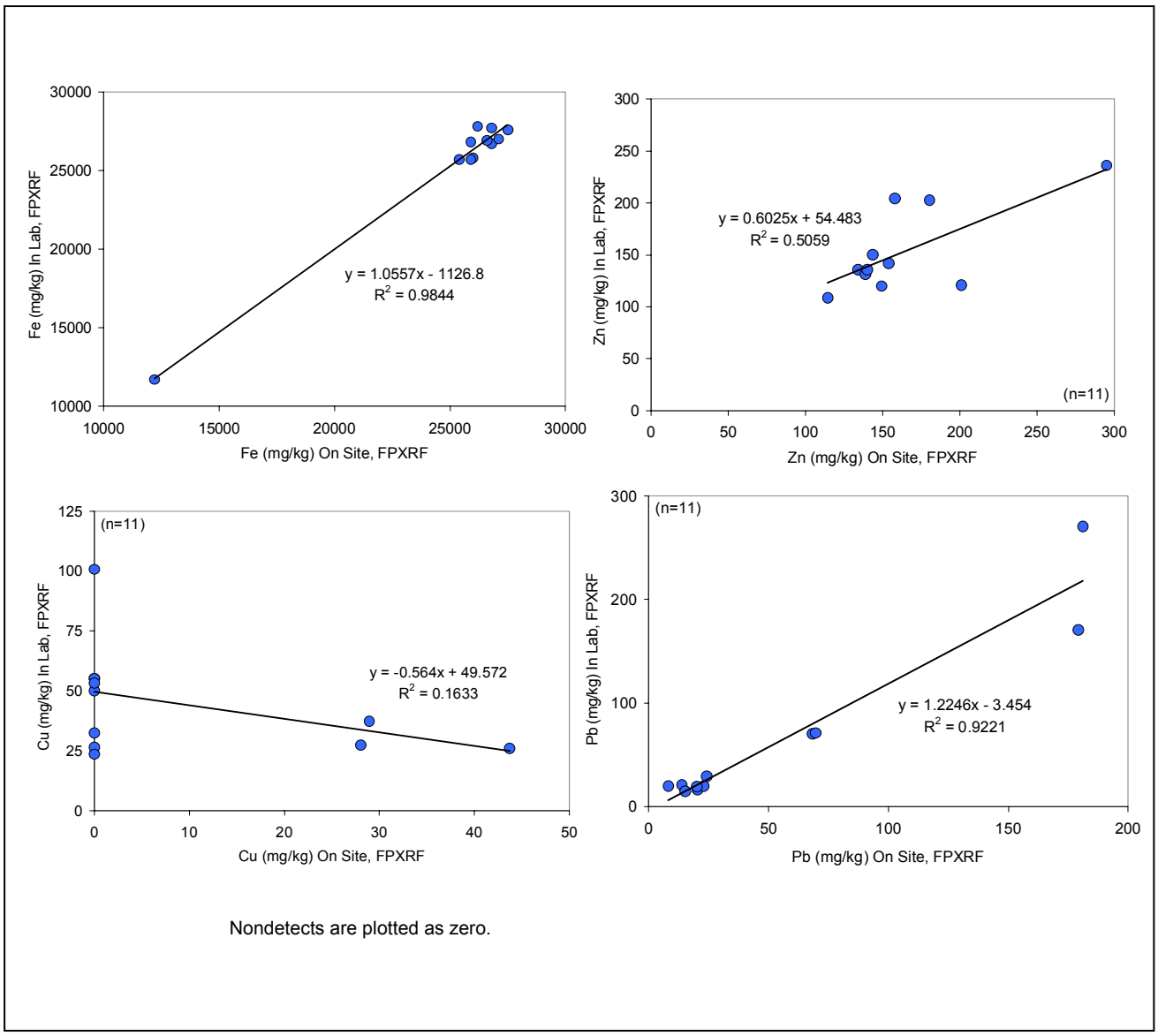


Figure 10. FPXRF results for Fe, Zn, Cu, and Pb (mg/kg wet) measured on site (Seaplane Lagoon) and in the laboratory (n=11).

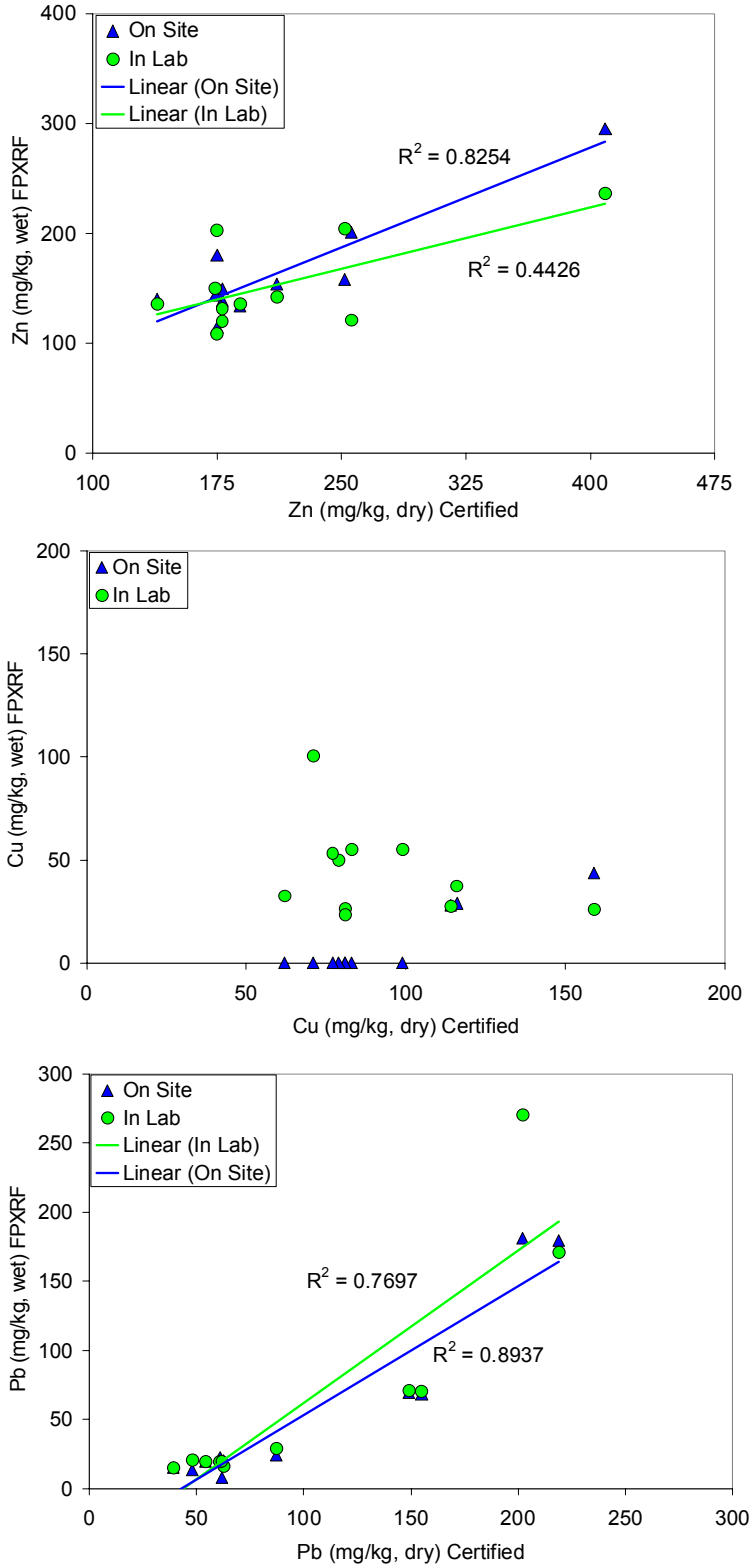


Figure 11. FPXRF results for Zn, Cu, and Pb measured on-site (Seapoint Lagoon) and in-laboratory (mg/kg wet), plotted verses certified results (mg/kg dry)..

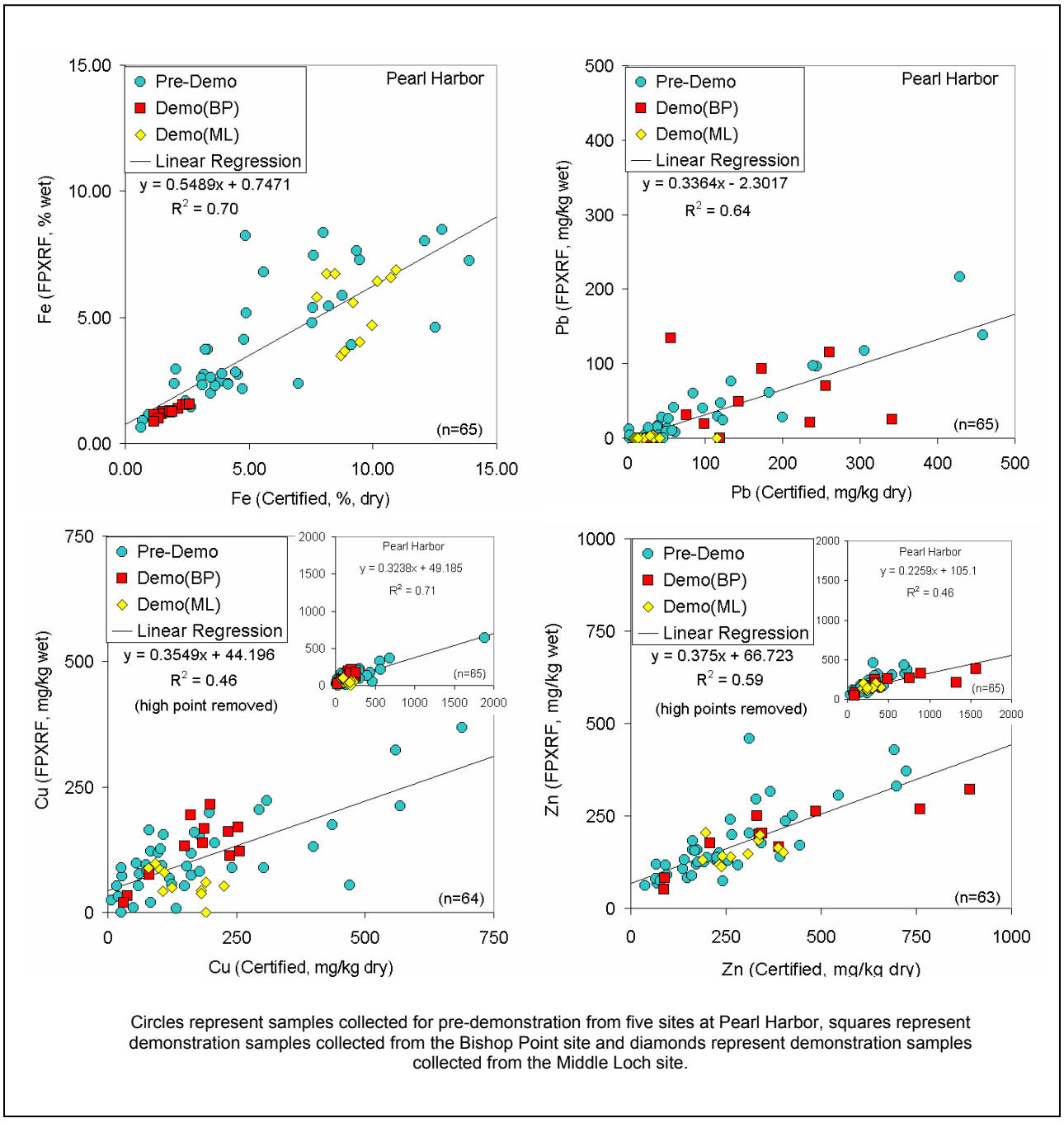


Figure 12. FPXRF Fe, Zn, Cu, and Pb results plotted against results from standard methods.

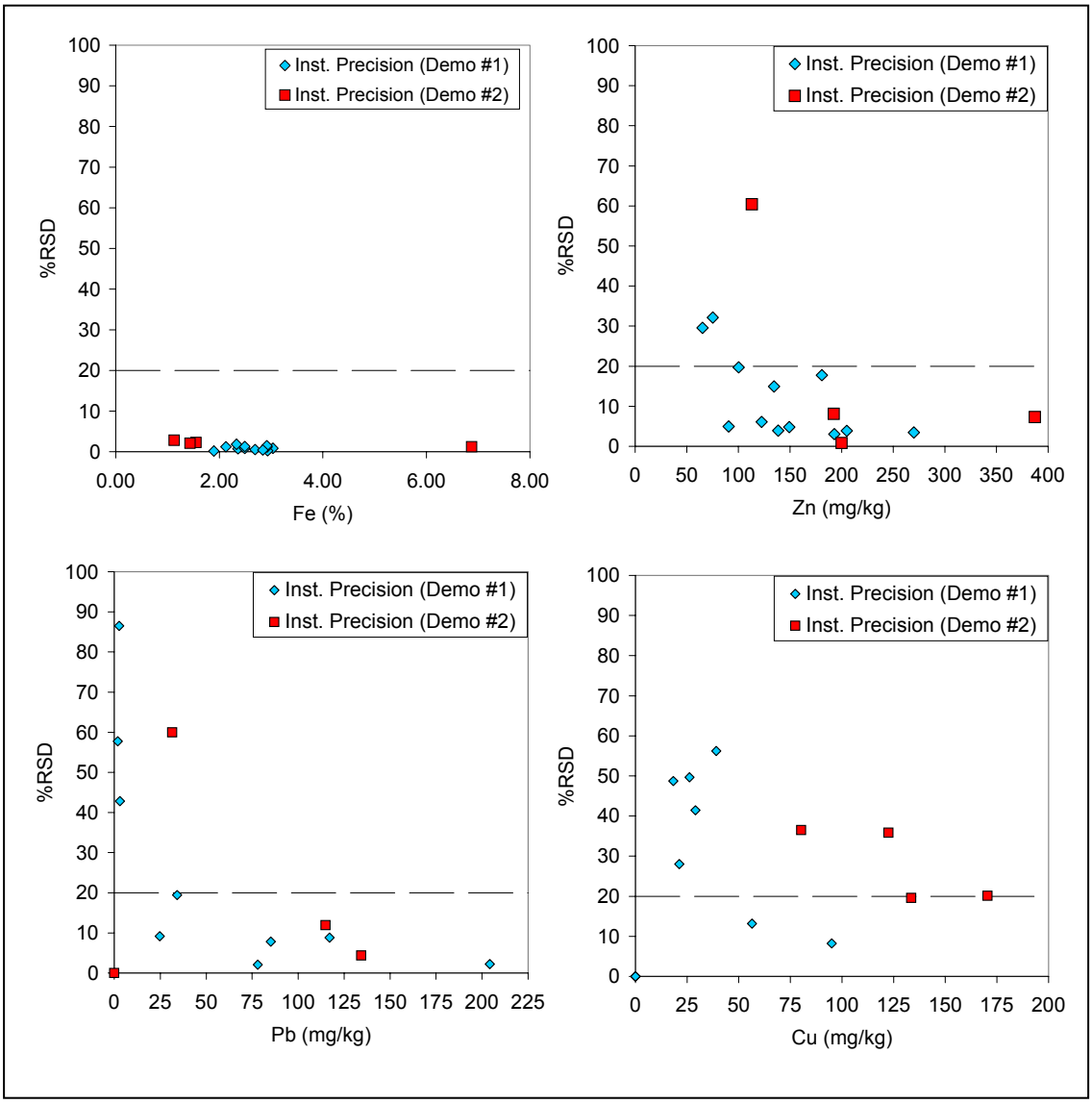


Figure 13. Instrument precision measurements for FPXRF (Fe, Zn, Pb and Cu).

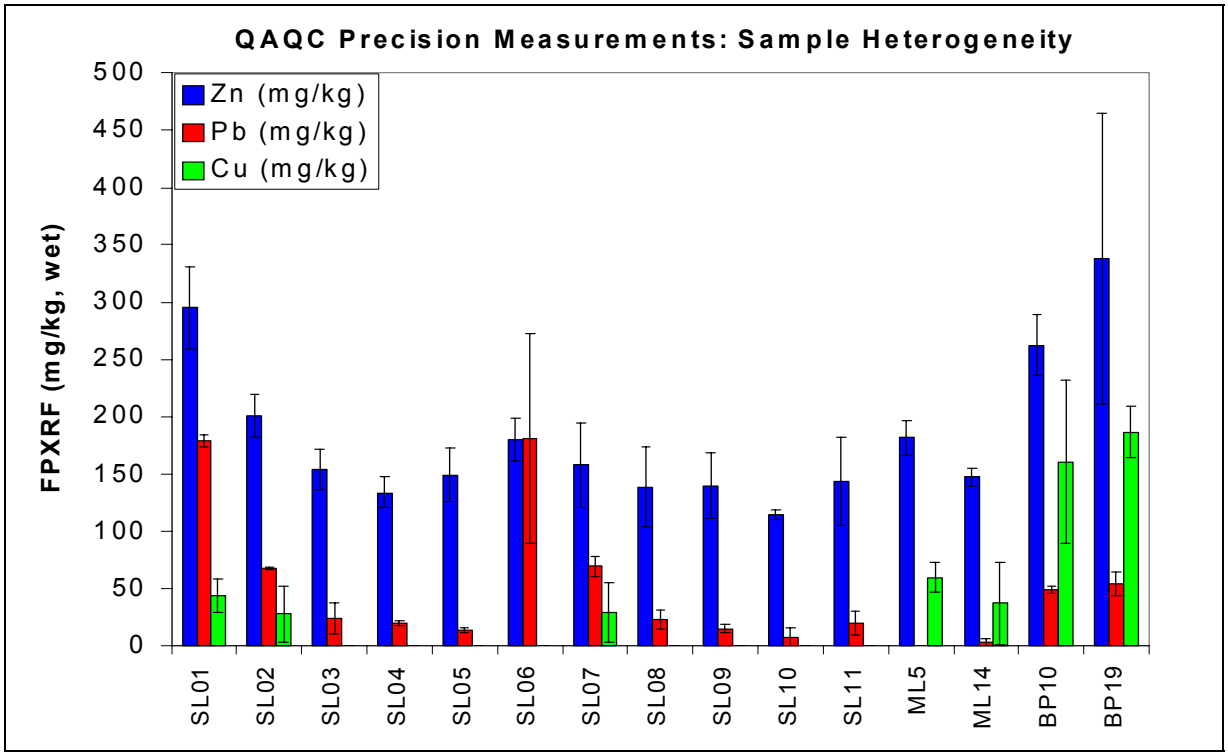


Figure 14. Combined results from heterogeneity-dependant precision analyses for Zn, Pb and Cu from Demonstration #1 and #2.

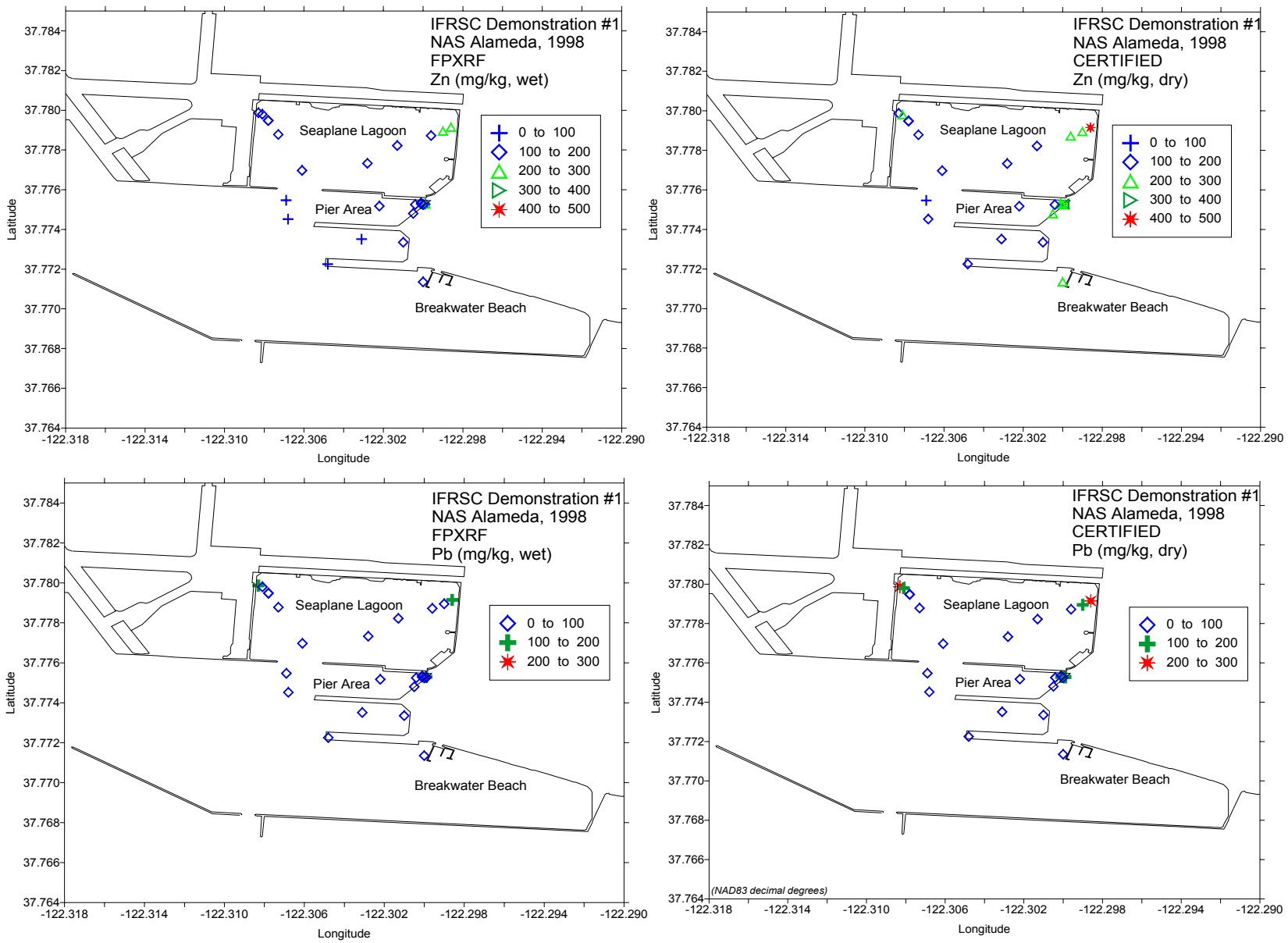


Figure 15. Comparison of maps using FPXRF results (left) and Certified results (right) for Zn and Pb at NAS Alameda (Demo #1).

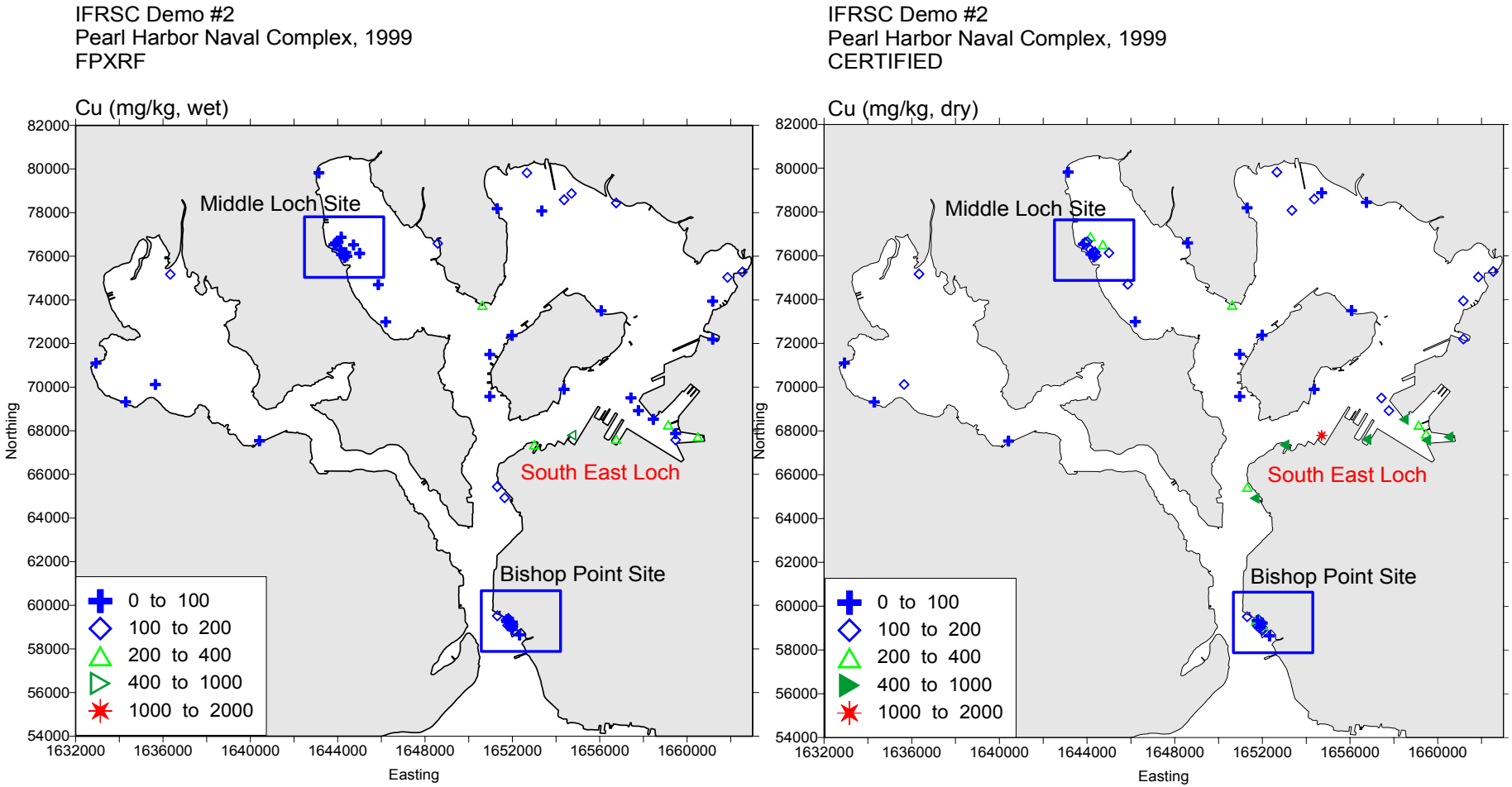


Figure 16. Comparison of maps using FPXRF results (left) and Certified results (right) for Cu at Pearl Harbor Naval Complex (Demo #2).

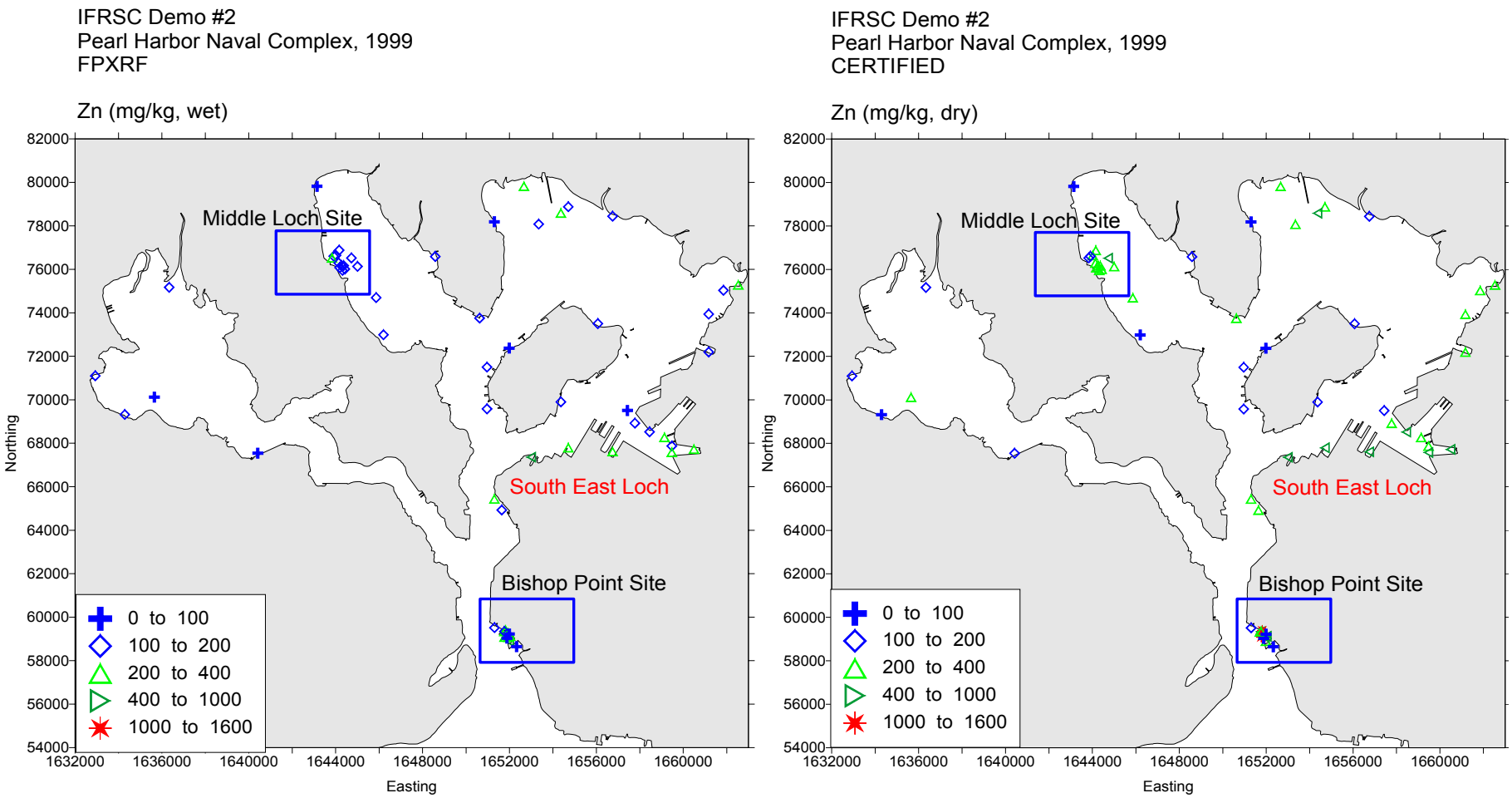


Figure 17. Comparison of maps using FPXRF results (left) and Certified results (right) for Zn at Pearl Harbor Naval Complex (Demo #2).

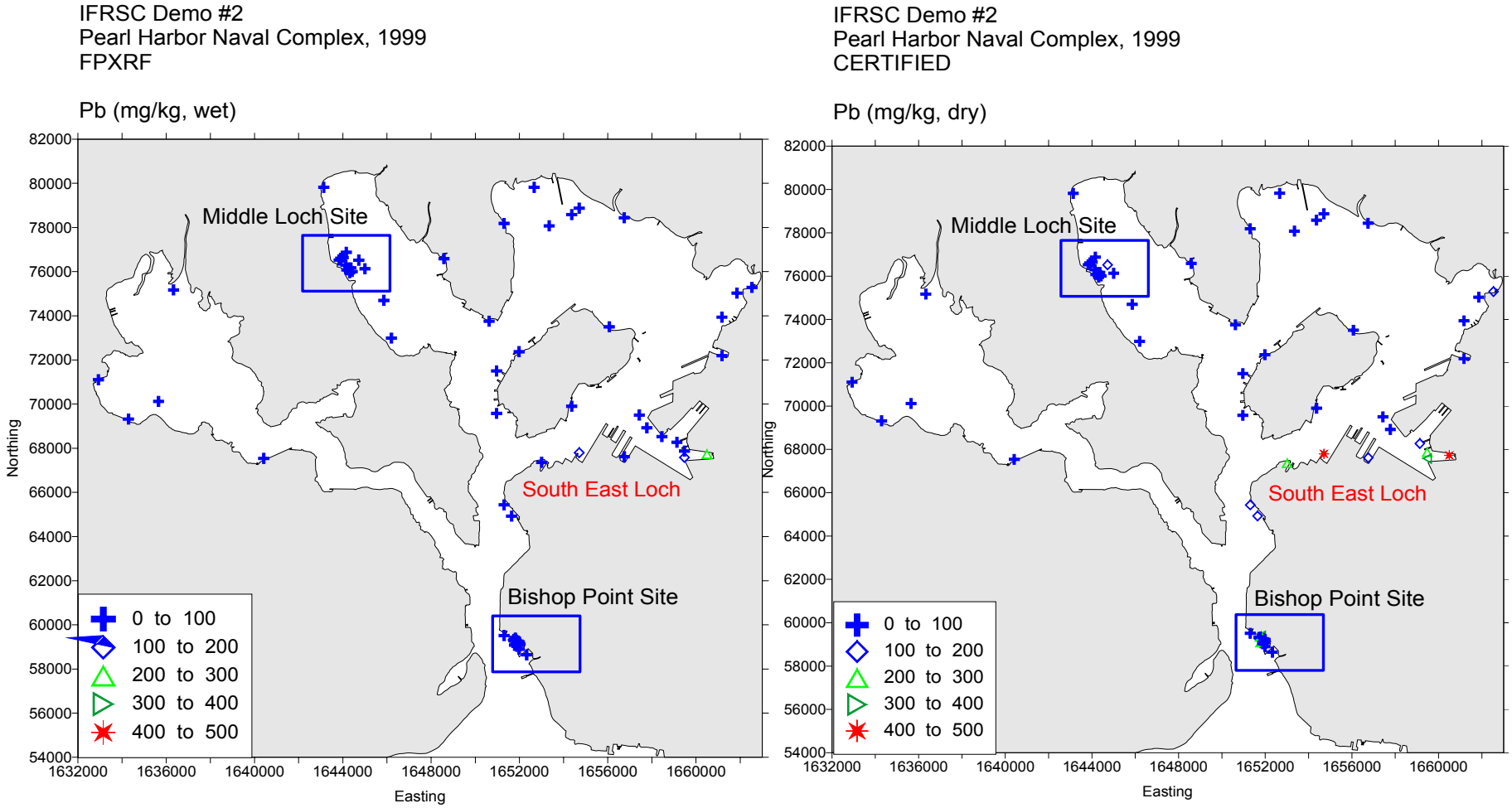


Figure 18. Comparison of maps using FPXRF results (left) and Certified results (right) for Pb at Pearl Harbor Naval Complex (Demo #2).

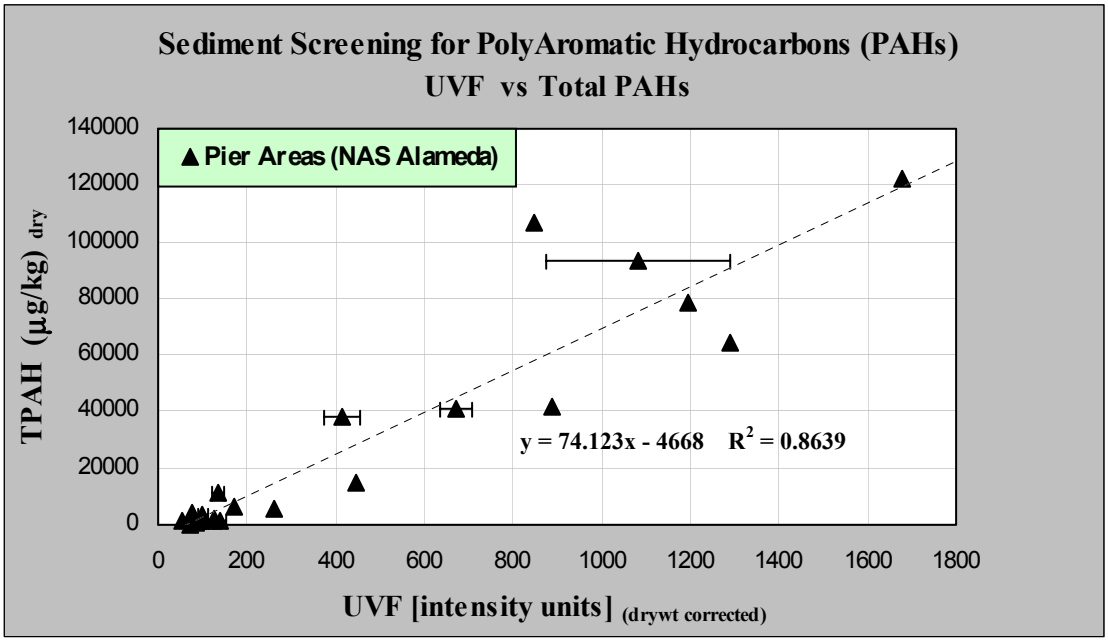


Figure 19. Predemonstration and Demonstration samples at Pier Area in Demo #1.

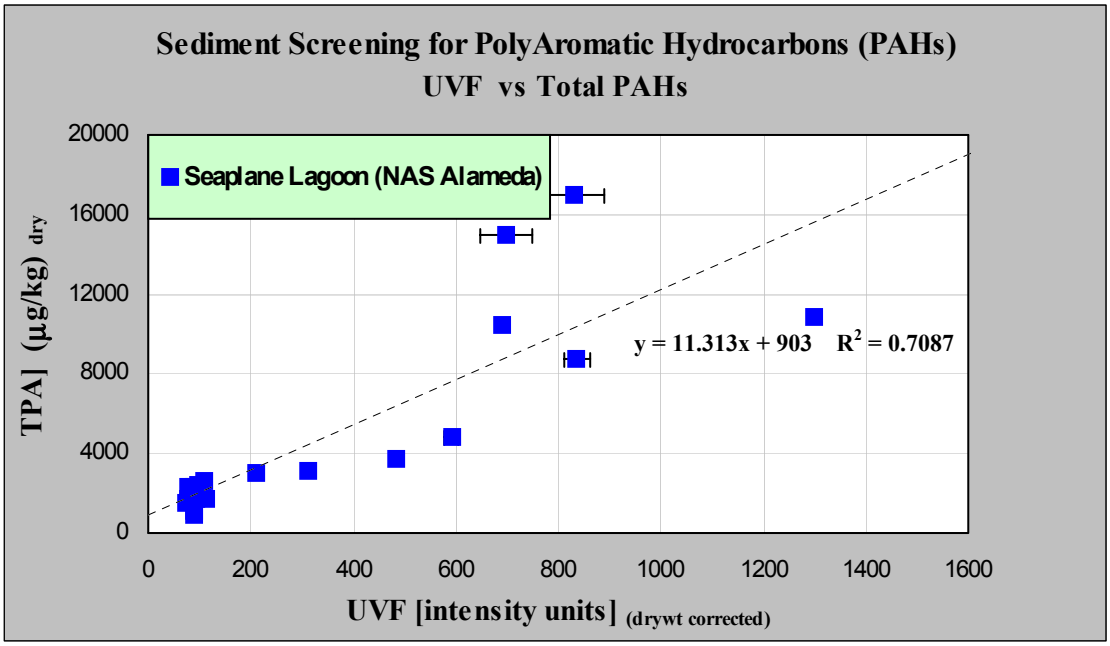


Figure 20. Predemonstration and Demonstration samples at Seaplane Lagoon in Demo#1.

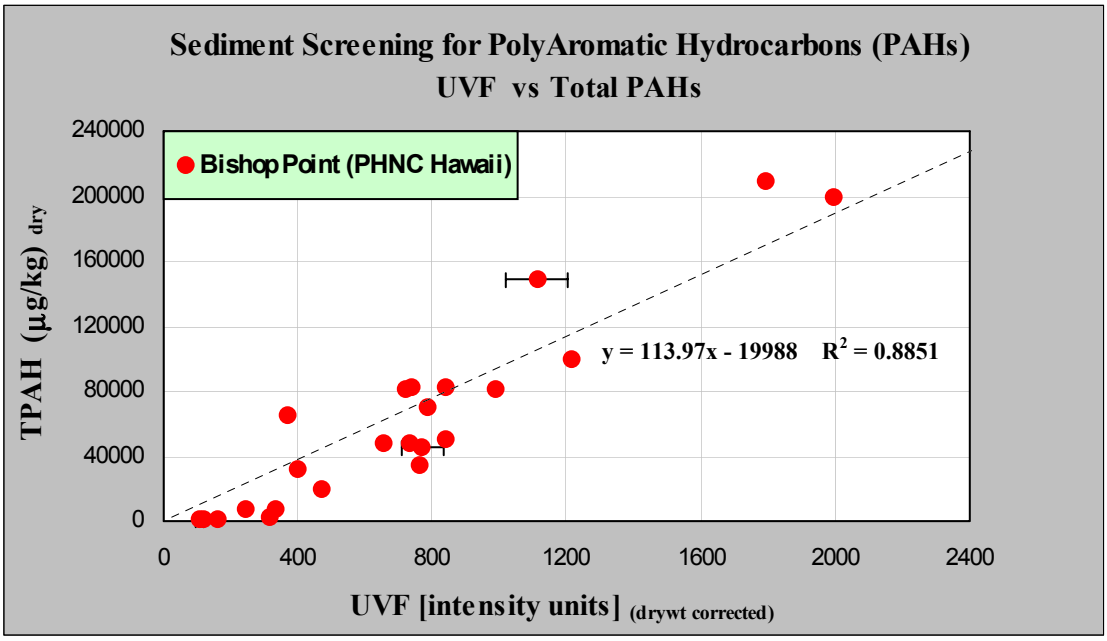


Figure 21. Predemonstration and Demonstration samples at Demo #2.

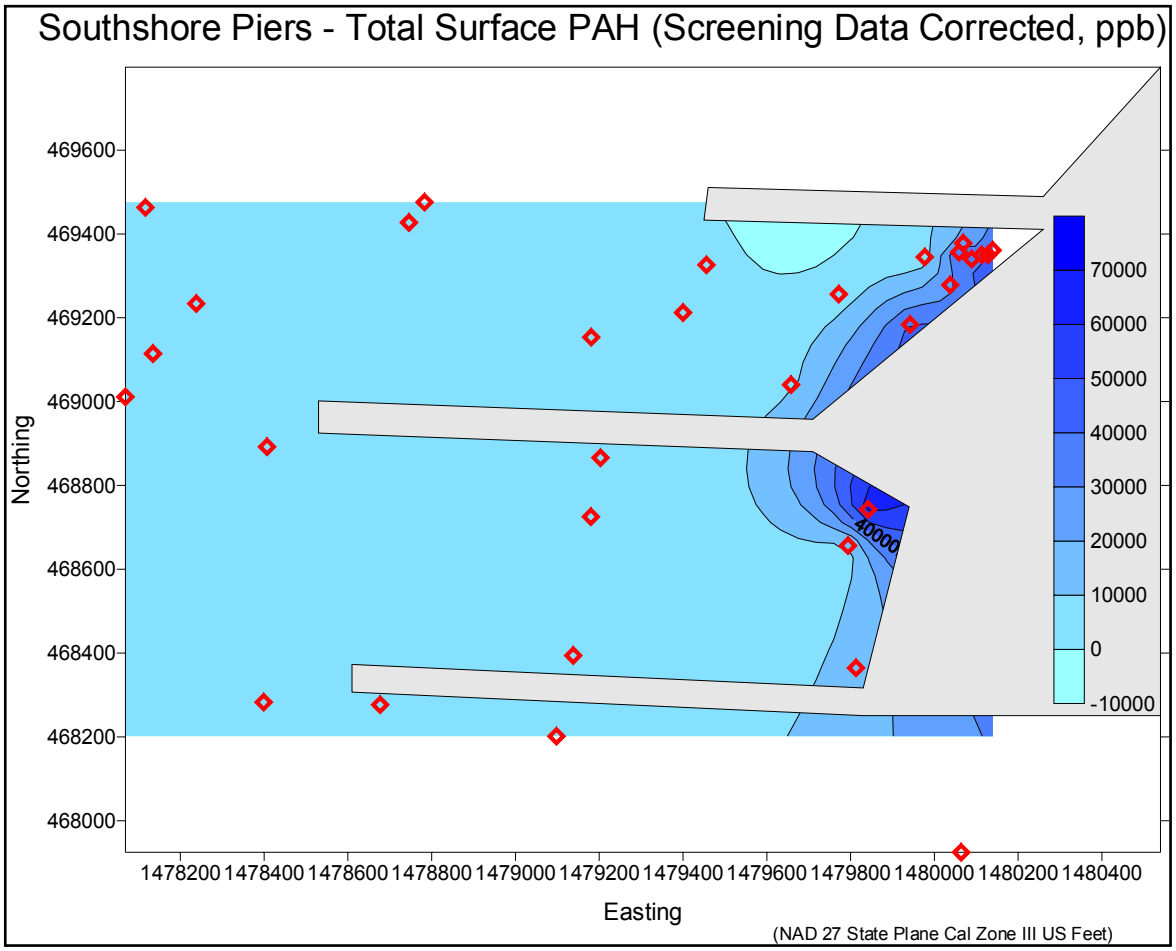


Figure 19. Contour map example showing spatial relationships in Pier Area data.

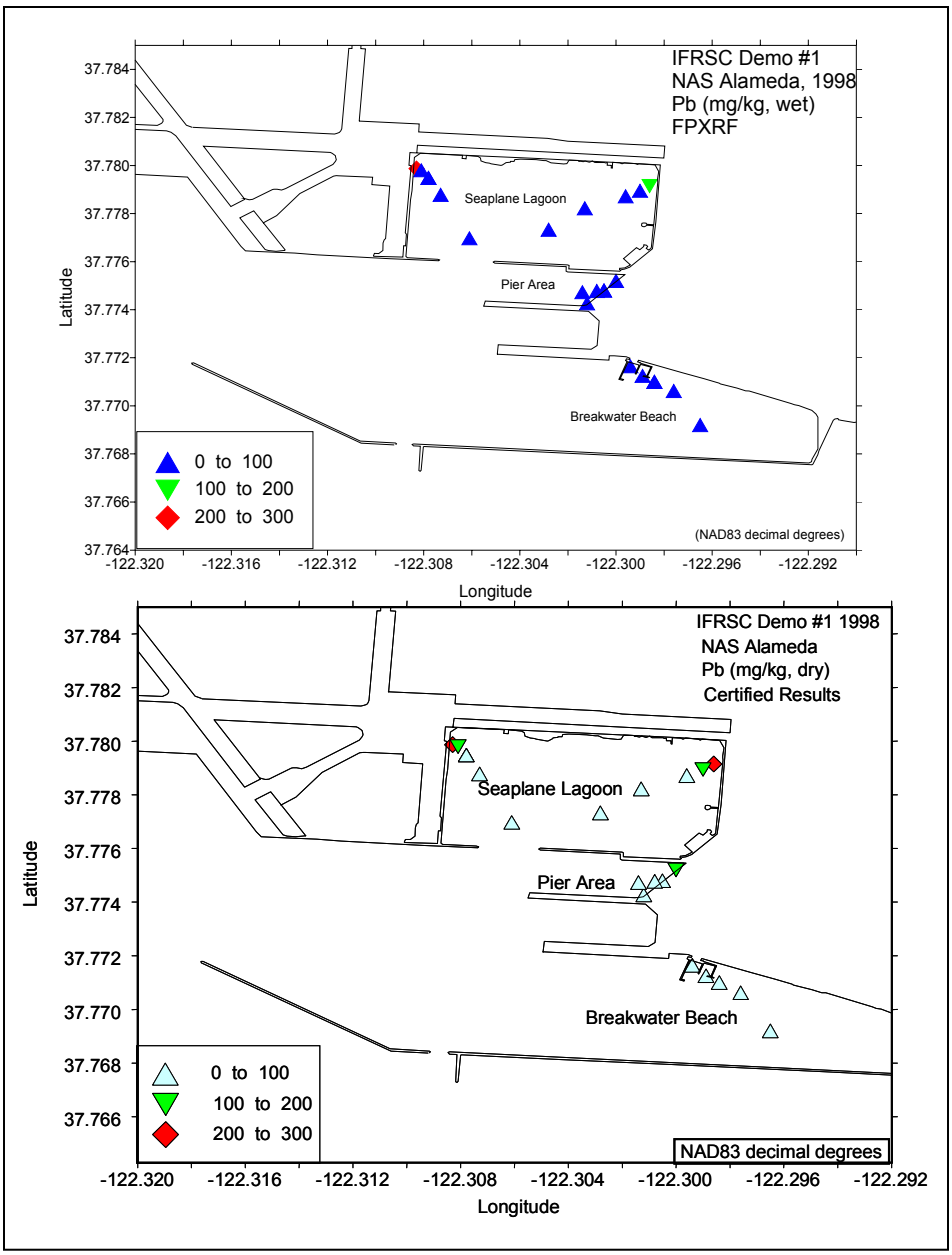


Figure 20. Laboratory results for tPAH, NAS Alameda.

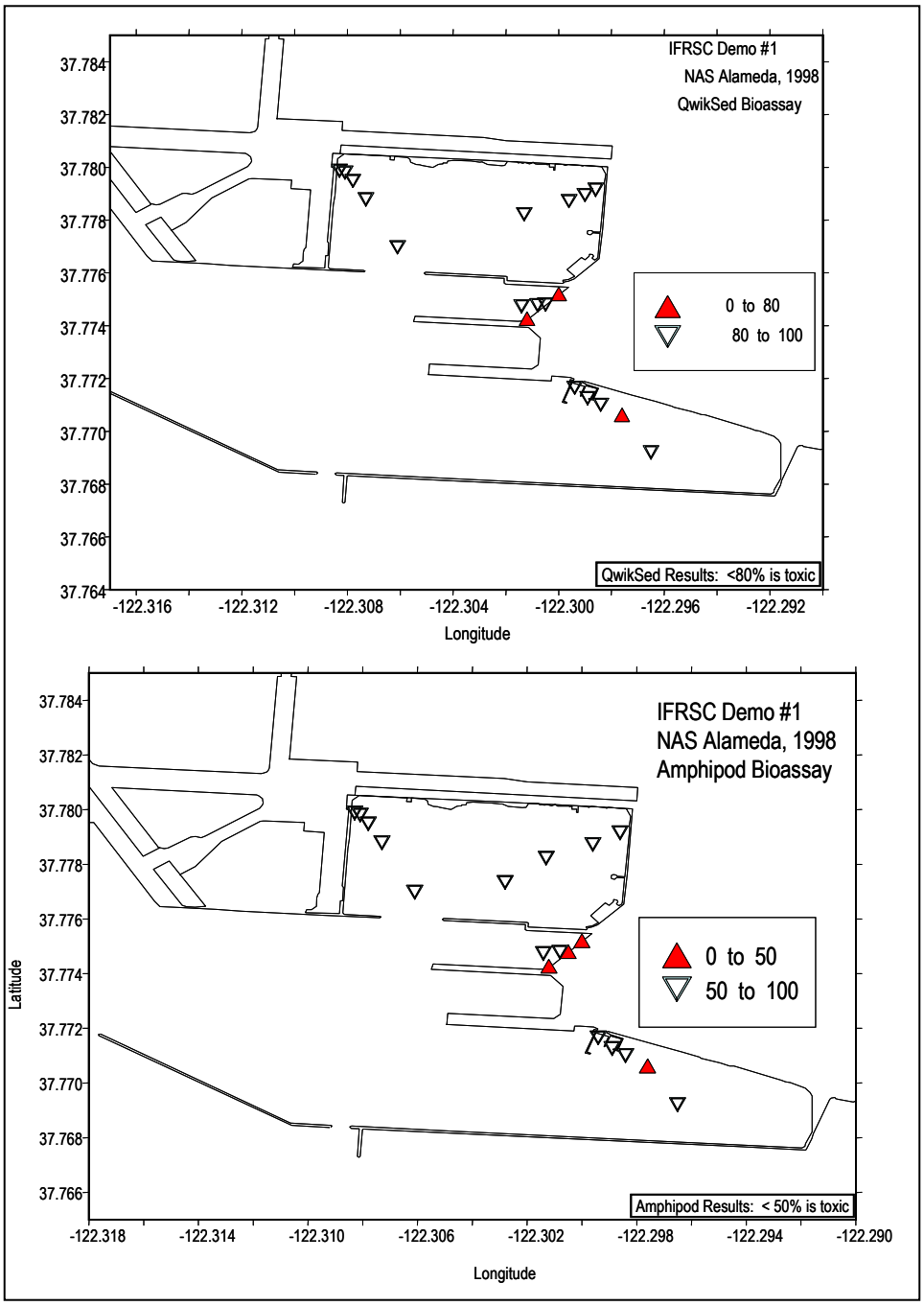


Figure 21. QwikSed and laboratory sediment toxicity results, NAS Alameda.

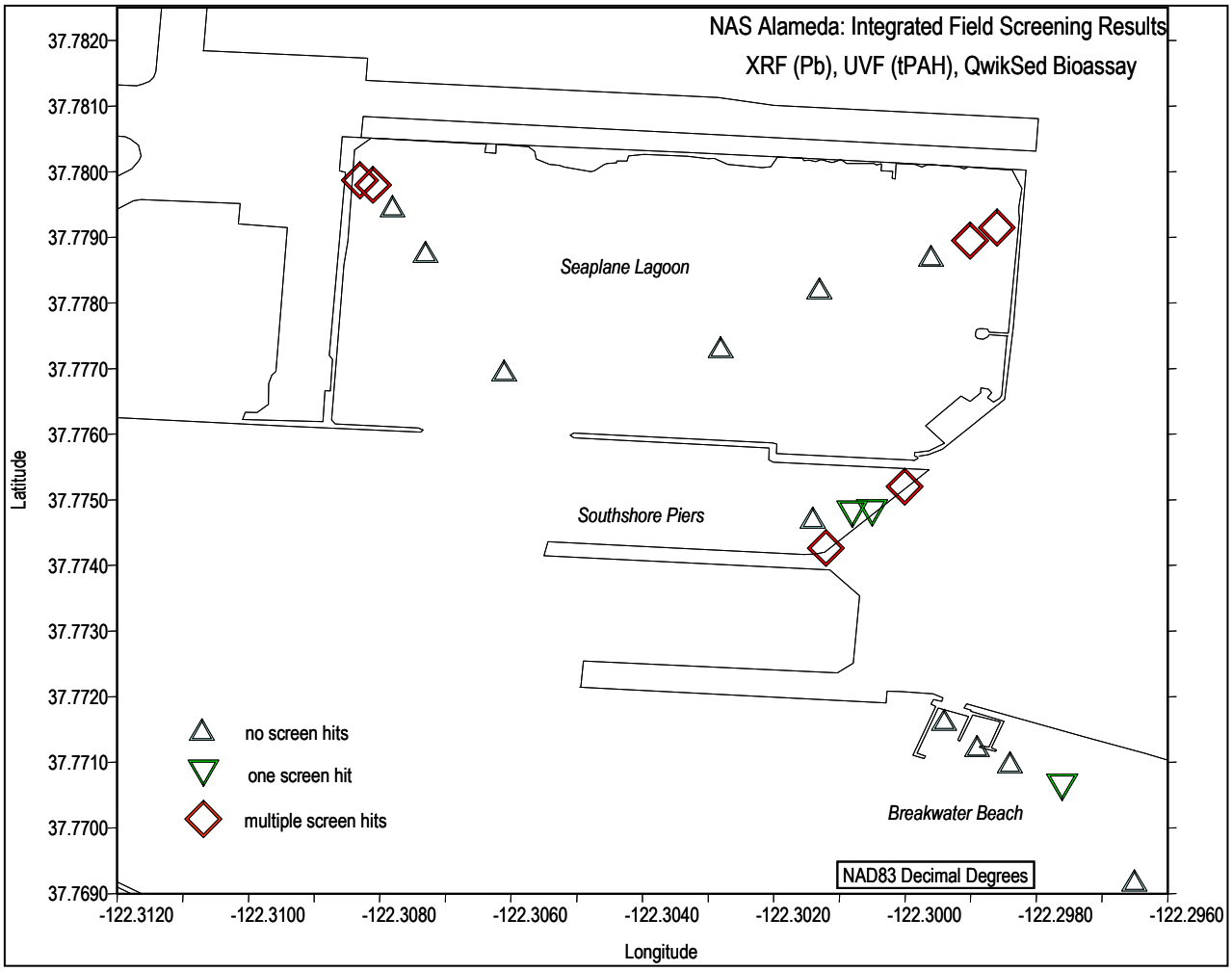


Figure 22. Integrated field screening results for NAS Alameda.

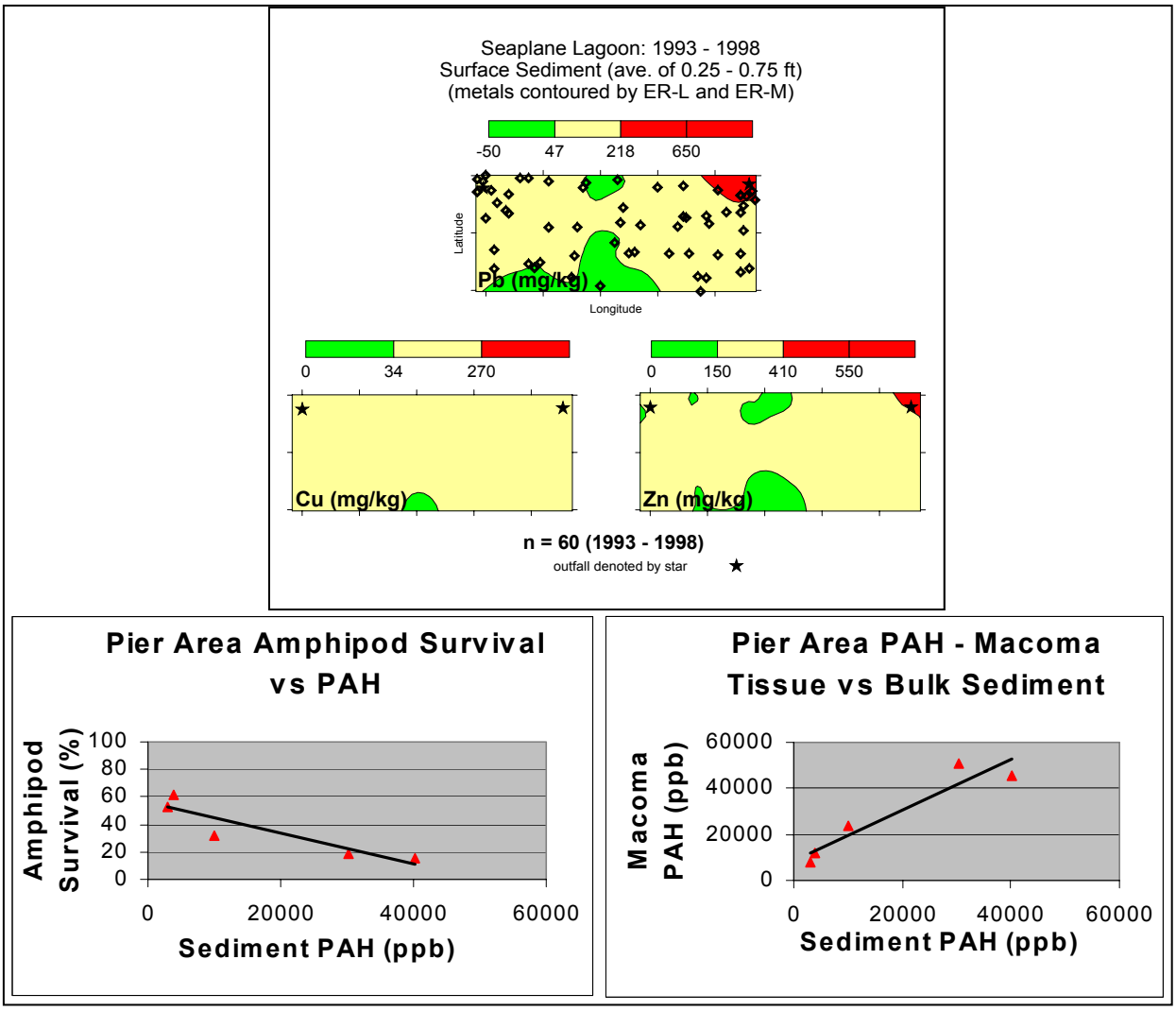

Seaplane Lagoon: 1993 - 1998
Surface Sediment (ave. of 0.25 - 0.75 ft)
(metals contoured by ER-L and ER-M)
-50
47
218
650
Latitude
Pb (mg/kg)
Longitude
0
34
270
Cu (mg/kg)
0
150
410
550
Zn (mg/kg)
n = 60 (1993 - 1998)
outfall denoted by star
Pier Area Amphipod Survival vs PAH
Amphipod Survival (%)
100
80
60
40
20
0
0
20000
40000
60000
Sediment PAH (ppb)
Pier Area PAH - Macoma Tissue vs Bulk Sediment
Macoma PAH (ppb)
60000
40000
20000
0
0
20000
40000
60000
Sediment PAH (ppb)

Figure 23.

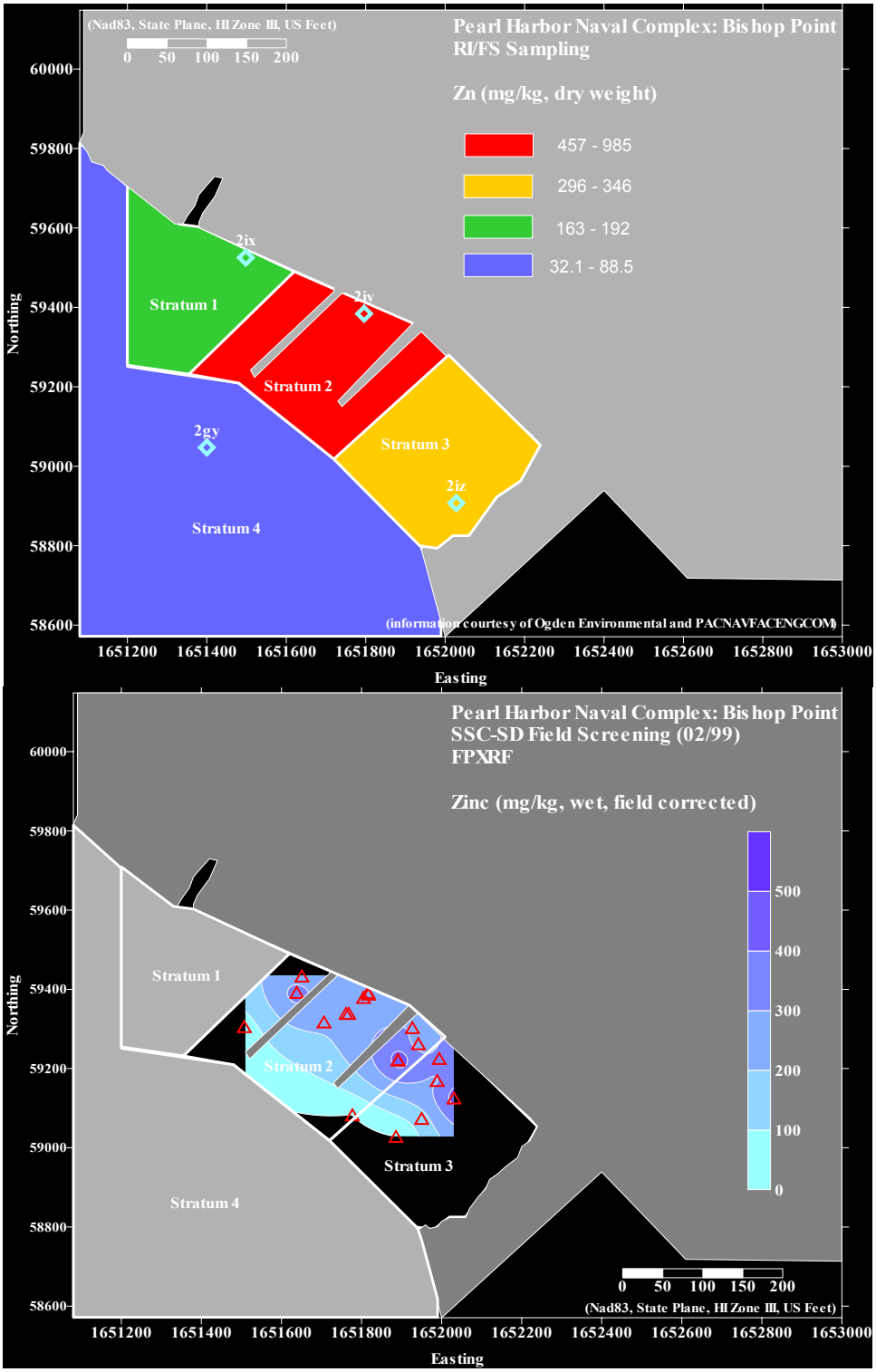


Figure 24. XRF results for zinc, Bishop Point, Pearl Harbor, HI.

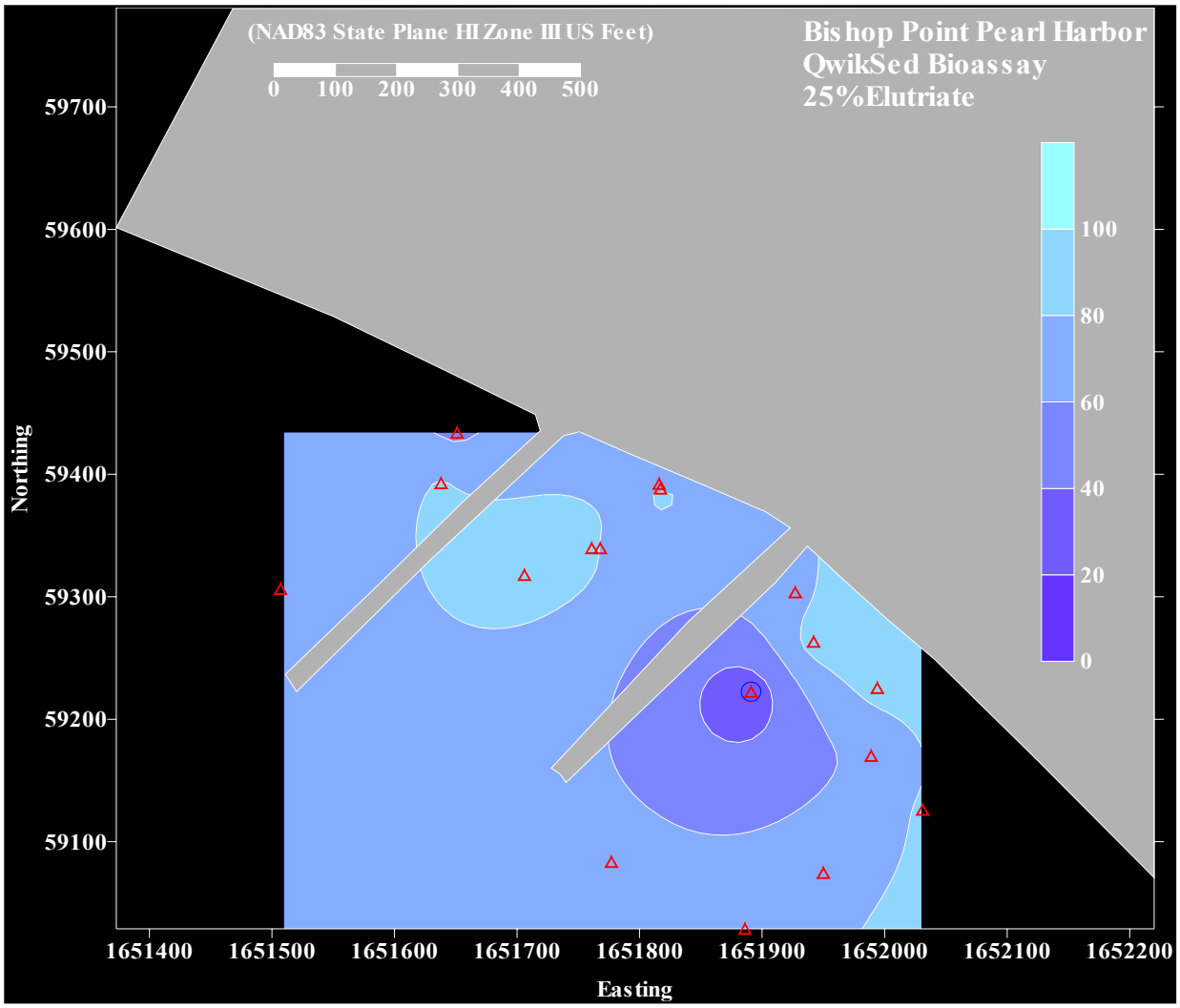


Figure 25. QwikSed sediment toxicity results for Bishop Point, Pearl Harbor, HI.

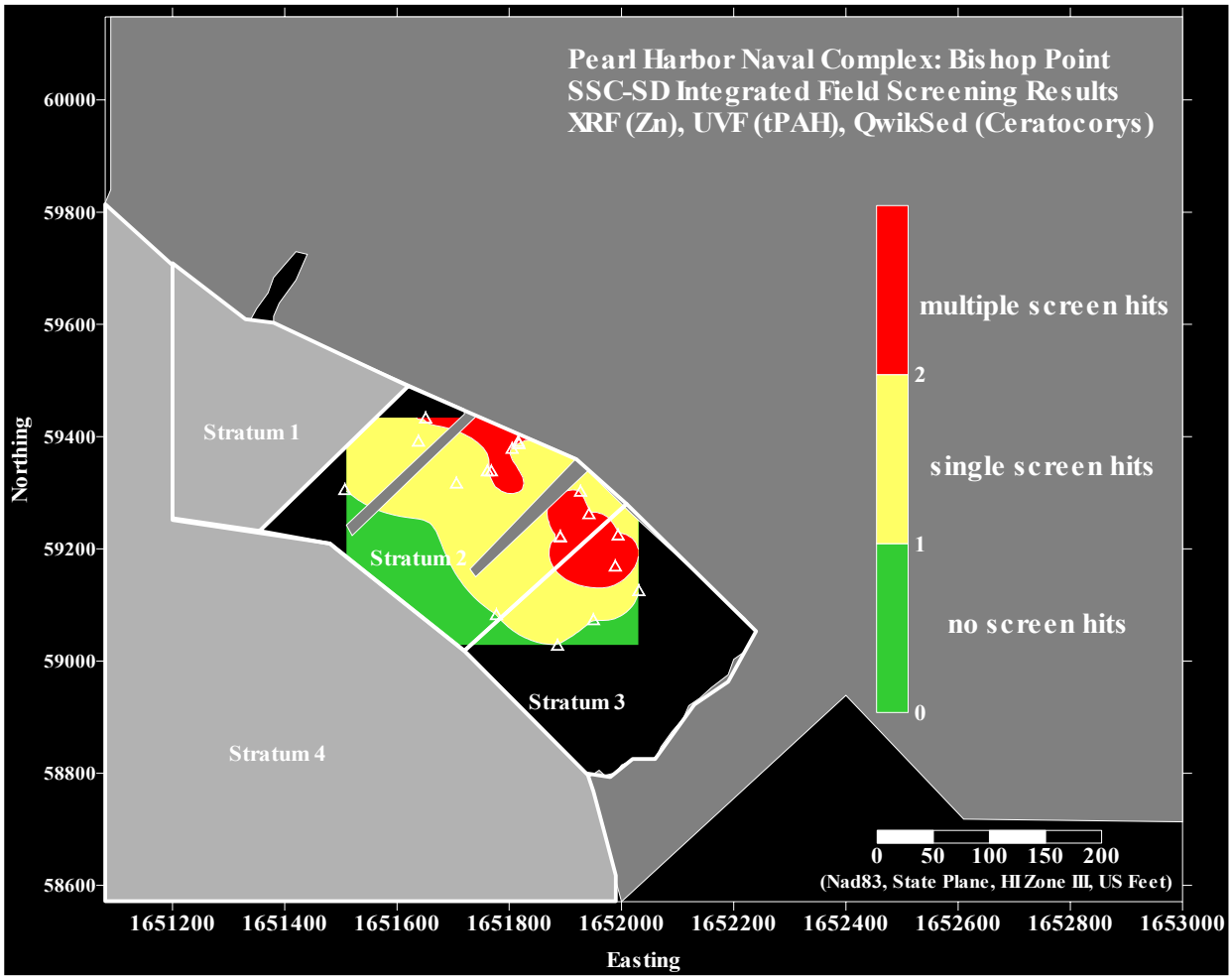


Figure 26. Integrated field screening results for Bishop Point.

Appendix A

Points of Contact

Appendix B

Table 1. QwikSed Data – Pearl Harbor, Hawaii

*Note: QwikSed tests using *Ceratocorys horrida*; 7-12 repeats

Sample ID	% of Control (100% elutriate)	% of Control (50% elutriate)	% of Control (25% elutriate)	IC50 (%)	Total Ammonia (ppm)	Unionized Ammonia (ppm	pH	Temp (C)	Salinity (ppt)	Date Collected	Date Tested
ML1	73	86.23	78.78	>100	0.56	0.02504	8.06	23	33	2/2/99	2/4/99
ML2	67.36	62.39	82.63	>100	0.162	0.00645	8.03	22.3	33	2/2/99	2/4/99
ML3	81.29	93.36	97.44	>100	0.0767	0.00279	7.99	22.3	33	2/2/99	2/4/99
ML6	84.06	101.22	94.78	>100	0.621	0.05130	8.27	25.6	34	2/2/99	2/3/99
ML7	107.99	98.34	93.54	>100	0.618	0.03317	8.15	22.8	32	2/2/99	2/5/99
ML8	90.27	92.05	94.27	>100	0.263	0.00722	7.85	22.7	33	2/2/99	2/5/99
ML9	118.72	103.03	94.49	>100	0.118	0.00423	7.96	23	32	2/2/99	2/5/99
ML10	143.01	146.56	126.78	>100	0.32	0.01004	7.92	22.7	34	2/2/99	2/5/99
ML11	77.49	70.35	56.46	>100	0.163	0.00508	7.92	22.6	34	2/2/99	2/5/99
ML13	73.28	93.53	99	>100	0.289	0.02193	8.23	25.6	34	2/2/99	2/3/99
ML14	86.12	77.30	70.92	>100	0.467	0.03096	8.17	25.5	34	2/2/99	2/3/99

ML15	84.55	81.76	70.74	>100	0.303	0.01009	7.94	22.9	34	2/2/99	2/5/99
BP4	37.94	40.84	55	42.256	0.164	0.00859	8.1	24.3	34	2/8/99	2/9/99
BP5	56.82	67.75	81.68	>100	0.186	0.01025	8.13	24.1	35	2/8/99	2/9/99
BP6	49.75	54.18	65.69	99.504	0.237	0.01288	8.12	23.9	33	2/8/99	2/9/99
BP7	42.32	67.28	83.63	84.62	0.315	0.01744	8.09	25.4	35	2/8/99	2/10/99
BP7	38.38			84.62	0.315	0.01744	8.09	25.4	35	2/8/99	2/8/99
BP8	54.30	80.15	91.49	>99	0.178	0.00809	8	25.4	35	2/8/99	2/10/99
BP8	55.5			>100	0.178	0.00809	8	25.4	35	2/8/99	2/8/99
BP9	39.09	70.46	86.46	82.61	0.0723	0.00301	7.97	25.1	35	2/8/99	2/10/99
BP9	33.29			82.61	0.0723	0.00301	7.97	25.1	35	2/8/99	2/8/99
BP10	54.53	71.89	89.25	>99	0.0846	0.00337	7.95	25.1	35	2/8/99	2/10/99
BP10	67.96			>100	0.0846	0.00337	7.95	25.1	35	2/8/99	2/8/99
BP11	15.43	23.24	26.12	10.55	0.0761	0.00301	7.94	25.3	35	2/8/99	2/10/99
BP11	20.74			10.55	0.0761	0.00301	7.94	25.3	35	2/8/99	2/8/99
BP12	37.59	64.61	70.05	77.04	0.107	0.00423	7.93	25.6	35	2/8/99	2/10/99
BP12	47.23			77.04	0.107	0.00423	7.93	25.6	35	2/8/99	2/8/99
BP13	63.25	68.45	82.64	>100	0.186	0.00975	8.11	24	35	2/8/99	2/9/99

BP14	43.78	59.79	73.98	88.937	0.182	0.00874	8.07	24	35	2/8/99	2/9/99
BP15	60.13	61.15	73.92	>100	0.195	0.00876	8.04	24	35	2/8/99	2/9/99
BP17	49.98	58.17	69.25	99.953	0.265	0.00934	7.93	24	35	2/9/99	2/10/99
BP19	57.67	62.53	72.67	>100	0.138	0.00662	8.07	24	35	2/9/99	2/10/99
BP20	59.18	69.58	73.46	>100	0.228	0.01144	8.1	23.7	35	2/9/99	2/10/99
BP21	65.11	74.20	87.19	>100	0.138	0.00606	8.01	24.6	35	2/9/99	2/10/99
BP22	49.62	59.70	63.34	98.849	0.144	0.00665	8.03	24.7	35	2/9/99	2/10/99
BP23	38.02	62.35	73.5	80.67	0.258	0.01495	8.14	24.5	35	2/9/99	2/10/99
NH3				2.839ppm							2/4/99
NH3				3.857ppm							2/5/99
NH3				3.828ppm							2/8/99
NH3				2.325ppm							2/9/99
NH3				1.908ppm							2/10/99
Cu				134ppb							2/4/99
Cu				>250ppb							2/8/99
Cu				129ppb							2/9/99
Cu				126ppb							2/10/99