ABSTRACT

Title of Thesis: TOXICITY AND CONTAMINATION IN BEAR CREEK SEDIMENT: SPATIAL ANALYSIS AND IMPLICATIONS FOR RISK ASSESSMENT

Sharon Hartzell, Master of Science, 2016

Thesis Directed By: Dr. Lance Yonkos
Department of Environmental Science & Technology

The sediments of Bear Creek near Baltimore, Maryland demonstrate substantial toxicity to benthic organisms, and contain a complex mixture of organic and inorganic contaminants. The present study maps the spatial extent and depth profile of toxicity and contamination in Bear Creek, and explores correlations between heavy metals, organic contaminants, and toxic responses. Two novel analytical techniques – handheld XRF and an antibody-based PAH biosensor – were applied to samples from the site to quantify total metals and total PAHs in sediments. By comprehensively assessing toxicity in Bear Creek, the present study provides data to inform future risk assessments and management decisions relating for the site, while demonstrating the benefits of applying joint biological assays and chemical assessment methods to sediments with complex contaminant mixtures.

TOXICITY AND CONTAMINATION IN BEAR CREEK SEDIMENT: SPATIAL ANALYSIS AND IMPLICATIONS FOR RISK ASSESSMENT

by

Sharon Elizabeth Hartzell

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2016

Advisory Committee:
Dr. Lance Yonkos, Chair
Dr. Michael A. Unger
Dr. Beth McGee
Dr. Sacoby Wilson

Acknowledgements

Thank you to my advisor and committee chair, Dr. Lance Yonkos, for his continued support throughout my master's thesis research. Dr. Yonkos has served as an endless well of knowledge, guidance and enthusiasm throughout the project. Thank you also to my committee members, Dr. Mike Unger, Dr. Beth McGee, and Dr. Sacoby Wilson for their generosity, insight and unique contributions to shaping this project.

Throughout this process, I have been lucky to receive support from a number of individuals who assisted with sample collection and preparation and contributed long days to ensure that toxicity tests were run successfully. Thank you to Elizabeth Friedel, Alex MacLeod, Allison Satterfield, Amy Wherry, Wenqi Hou, Barret Wessel, Gary Seibel, Dr. Ray Weil, and Dr. Daniel Fisher, who all contributed to various aspects of the project at College Park, and to Mary Ann Vogelbein and George Vadas at VIMS who instructed me in PAH biosensor analysis and GC-MS sample preparation. Thank you to Lee Drake of Bruker, for his help with XRF analysis. Thank you also to the faculty and fellow graduate students of the Department of Environmental Science and Technology for their camaraderie and encouragement.

Lastly, I would like to thank my friends and family for support throughout the graduate school process. In particular, I thank Chip and Daryl Hartzell, for their constant encouragement; Annette Hartzell, for her support and empathy; and Brittany Reynoso, for her love and patience over the past two years.

***

This project would not have been possible without support for my graduate education from the University of Maryland Department of Environmental Science & Technology. Funding for toxicity assays was provided by the Chesapeake Bay Foundation, and funding for PAH biosensor analysis was provided through the National Institute of Environmental Health Sciences – Superfund Research Program grant number 5RO1ES020949-03.

Table of Contents

List of Tables

List of Figures

Chapter 1: Spatial Assessment of Toxicity and Contamination in Bear Creek

INTRODUCTION

Site overview

Contaminated sediments present environmental and public health concerns in the United States and worldwide. A 2002 assessment of coastal estuaries in the United States determined that 35% of estuaries are degraded due to contaminated sediments (United States EPA (USEPA), 2002). According to the USEPA, contaminated sediments pose ecological and/or human health risks in many regions throughout the country, even in areas where water quality criteria are not exceeded, and that harbors are particularly common hot spots for pollution (USEPA, 2004). Contaminated sediments "are typically the ultimate repository for contaminants in the environment," resulting from both point and nonpoint pollutant sources; therefore, they serve as reservoirs and potential long-term sources of contaminants back into waterways and the environment (Reible and Lanczos, 2006). Contaminated sediments pose threats to ecological integrity and human health, while reducing the recreational use of rivers and harbors in the U.S., and contributing to the enactment of fish consumption advisories (USEPA, 2005). Remediation activities are often undertaken to prevent risks posed by contaminated sediments, but site-specific understanding of contaminant extent and transport is necessary to decide the best course of remedial action and environmental management (USEPA, 2005).

The Baltimore Harbor-Patapsco River system (Maryland, USA) is one of the most contaminated areas of the Chesapeake Bay. It has been designated as one of three Regions of Concern in the Bay by the Chesapeake Executive Council, and, along with the Anacostia and Elizabeth Rivers, has been the focus of USEPA study for over two decades (Chesapeake Bay Program, 1997). By 2007, Baltimore Harbor was still identified as one of the most polluted areas in the country (Hartwell and Hamadi, 2007).

Figure 1.1. Aerial imagery of the Chesapeake Bay and Baltimore Harbor. Images from Environmental Systems Research Institute (ESRI) Basemap database

Baltimore Harbor has been a center of industrial production since the beginning of the 18th century (Mason et al., 2004). Iron-ore related activity in the region of the harbor began in the 1700s, with production increasing during the Revolutionary War (Mason et al., 2004). In the 1800s railroads were introduced and the area became a center of industrial activity, with coal as a major export (Mason et al., 2004). Industrial activity

expanded prior to World War II, with Baltimore factories supplying much of the demand for steel (Mason et al. 2004). Though environmental concerns were not a high priority throughout most of this history, the effects became known eventually; by the 1960s, the Patapsco River was intensely polluted by both industrial activities and sewage discharge (Mason et al., 2004).

Concerns for Baltimore Harbor have focused on the elevated concentrations of heavy metals and organic contaminants in sediments. In 1997, a comprehensive study was conducted to map concentrations of contaminants in surface sediments within the Baltimore Harbor/Patapsco River/Back River system. Samples were collected from 80 sites in the Patapsco River, Baltimore Harbor, Curtis Creek, the Back River, and smaller tributaries including Rock Creek, Stony Creek, Bear Creek, Colgate Creek, and Old Road Bay (Baker et al., 1997). The study analyzed these samples for polychlorinated biphenyls (PCBs), polycyclic aromatic hydrocarbons (PAHs,) organochlorine pesticides (OCs), and heavy metals including cadmium, chromium, copper, iron, lead, mercury, methylmercury, nickel, and zinc (Baker et al., 1997). Ashley and Baker (1999) determined that organic contaminants were highly variable throughout the system, and were not correlated with grain size or organic carbon content; they suggested that proximity to sources of input may have contributed to the variability. In general, lower PAH concentrations were observed in sediments with high sand content, which are located near to the shores of the channel (Ashley & Baker, 1999).

One region of particular concern was Bear Creek, a tributary of the Patapsco River located near Baltimore Harbor and directly adjacent to the former Sparrows Point steel manufacturing plant (Fig. 1.3). Bear Creek sediments were particularly elevated in

zinc, chromium, lead, hypothesized to be associated with a long history of steel production at the Sparrows Point complex adjacent to the creek (Baker et al., 1997). Sites in Bear Creek near Sparrows Point also showed the second highest concentration of total PAHs within the system, as well as considerable amounts of oil and tar which stymied accurate organic contaminant quantification (Baker et al., 1997). Bear Creek also had elevated levels of PCBs and organochlorine pesticides (Baker et al., 1997). Organic contaminants within the system were elevated in direct proximity to the Sparrow's Point Industrial Complex, likely associated with pyrolysis of coal during the production of steel (Ashley and Baker, 1999).

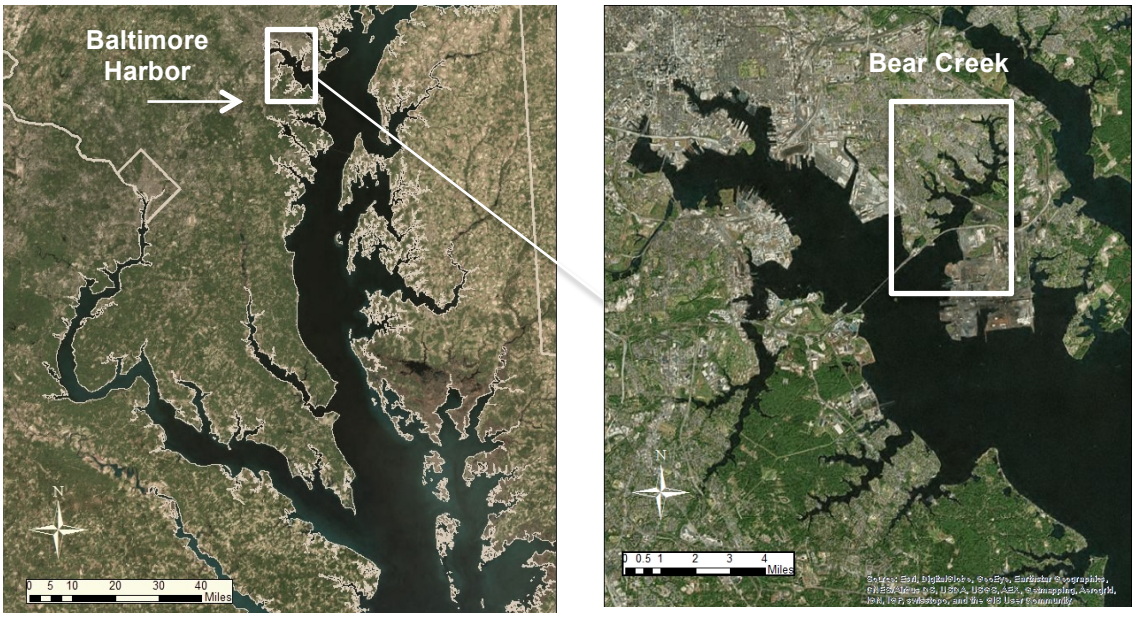


Figure 1.2. Aerial imagery of Baltimore Harbor/Bear Creek. Images from ESRI basemap database.

Figure 1.3. Aerial imagery of the Sparrows Point Industrial Complex/Bear Creek study site. Image from ESRI basemap database.

McGee et al. (1999) assessed a subsample of these sites for sediment toxicity using the 10-day acute tests with the estuarine amphipod *Leptocheirus plumulosus,* a recommended species for assessing marine and estuarine sediments (EPA, 2000). Two sites assessed in this study were located in Bear Creek, one of which yielded 0% survival and the other 50% survival (McGee et al., 1999). McGee et al. demonstrated continued toxicity with a subsequent test in 2004. In 2012, researchers at the Wye Research and Education Center demonstrated toxicity at five sites within Bear Creek, with 0% survival

in two sites within the offshore area and compromised survival in three other sites (Yonkos et al., 2012).

In 2006, a comprehensive study was conducted on Baltimore Harbor sediments, attempting to assess the causes of demonstrated toxicity in Bear Creek and other sites within Baltimore Harbor. Though contaminant classes and elevated toxicity were demonstrated, it had thus not been attempted to link toxicity to a particular chemical class, which, as they wrote, "greatly hinders management of the Harbor" (Klosterhaus & Baker, 2006). The study applied a whole-sediment toxicity identification and evaluation (TIE) approach to sediments from the site, but the authors were unable to distinguish classes of compounds potentially responsible for toxicity. They assessed both metals and organic compounds for their relationship to toxicity, but sediment manipulations that worked in the laboratory with spiked control sediments were unable to remove toxicity when applied to the sediments of Baltimore Harbor, likely due to their complexity and the presence of other compounds that are potentially responsible for toxic responses (Klosterhaus & Baker, 2006).

Other studies have explored the relationship between toxicity and various contaminant classes within the Baltimore Harbor system. In a 1999 study, McGee et al. posited that toxicity at stations within Bear Creek may have been driven by sediment-associated metals, while Inner Harbor toxicity could be connected with both metals and organic contaminants (McGee et al., 1999). They determined that organic compounds were not present at levels high enough to explain toxicity, though at one site directly adjacent to Sparrows Point, high levels of oil and tar interfered with analysis (McGee et al., 1999). Hansen et al. (1996) assessed Bear Creek sediments and found that metals-

related toxicity was not present in porewater, likely because of the high sulfide concentrations in the sediment. A comparison of Simultaneously Extracted Metals to Acid Volatile Sulfide (SEM/AVS) in sediments is often used to predict bioavailability and toxicity of contaminants, with ratios of less than one indicating low bioavailability and toxicity, as reactive sulfide binds metal species (Klosterhaus & Baker, 2006). In Bear Creek sediments, SEM/AVS ratios were all considerably less than one, indicating low bioavailability of metals (Klosterhaus & Baker, 2006). Graham et al. (2009) assessed chromium in sediment and porewater from 22 locations within the Baltimore Harbor system, and found low concentrations of Cr(VI), the more toxic species of the metal, in both sediments and sediment porewater.

Contamination at depth within Bear Creek and Baltimore Harbor has been explored to a limited degree. One report assessed a sediment core within the Bear Creek system, which showed consistent PAH concentrations over the past 30 years, and highest concentrations of metals (except zinc) within the surface layer (Klosterhaus & Baker, 2006). A study by Mason, Kim and Cornwell in 2004 also assessed metals concentrations at depth within sediment cores in Baltimore Harbor, relating contaminant concentrations to the industrial history of the area. One of their study sites was the historically toxic Site 28, located directly adjacent to Sparrows Point (Fig. 1.3), which was assessed and reported by McGee et al. (1999, 2004) and caused 100% mortality for benthic amphipods. This sediment core revealed peaks in chromium after 1960 with lower values at the surface; Zinc also peaked during the late 20th century with an even more substantial peak around 1920 (Mason et al., 2004). Lead and nickel were both lowest in concentration at the sediment surface, with the lead peak occurring around 1950 and the

nickel peak around 1980 (Mason et al., 2004). Researchers dated the core from Site 28 as a part of this study, and found that sediments in the area had been subject to repeated disturbances and variable sedimentation rates, but that the 2-meter core does not reflect any time before major human impacts in the area – likely representing a time period of 100 years (Mason et al., 2004).

Management status of Sparrows Point

Sparrows Point began steel production in the 1880s, and was under ownership by Bethlehem Steel for most of its history, until the company liquidated in 2003 (Reutter, 2012). The site changed hands several times, from Mittal Steel to Severstal to RG Steel, who ran the facility until they, too, went bankrupt in 2012 (Wood, 2016). After two years of ownership by Hilco Global, a financial services firm, it was bought in September, 2014 by Sparrows Point Terminal, LLC, which has entered into agreements with the USEPA and the Maryland Department of the Environment (MDE) to address pollution at the site. The firm has agreed to commit $48 million to environmental cleanup, as well as contributing funds to investigate issues in the offshore area of Sparrows Point (Wood, 2014). The company plans to redevelop the site into an industrial and transportation complex, and have renamed the property Tradepoint Atlantic (Wood, 2016). At the same time, the Maryland Port Authority has sought space to build a facility for dredged material storage at the site (Sherman, 2015). In light of planned redevelopment, a thorough understanding of the environmental and human health risks present at the site has become of paramount importance.

Ecological and human health risk assessments

Bear Creek, Sparrows Point, and the surrounding areas have been subject to a series of ecological and human health risk assessments within the past several years. In 2010, a Baseline Ecological Risk Assessment (ERA) was conducted for on-site areas at Sparrows Point. The assessment found elevated food chain risks present at the site for cadmium, chromium, lead, nickel, selenium, and zinc, as well as for high molecular weight PAHs, and concluded that soil invertebrates and benthic invertebrates were at risk from exposure, particularly to chromium (URS Corporation, 2010). In 2011, EA Engineering, Science and Technology, Inc. (EA) conducted a risk assessment of the offshore area of the Coke Point Peninsula, at the southwestern corner of the Sparrows Point complex (Figure 1.3), to assess potential risks to aquatic and benthic communities and to wildlife. The risk assessment was based on site measurements of contaminants in sediment and surface water, and comparison to toxicological benchmarks. They found sediment concentrations of arsenic, chromium, copper, lead, mercury, zinc, dioxins, high molecular weight PAHs, low molecular weight PAHs, and PCBs at levels that were likely to cause risks to wildlife (EA, 2011). For surface water, several metals, ethylbenzene, toluene, and PAHs exceeded benchmarks, with PAHs as the most likely to drive risks in near-shore areas (EA, 2011). In assessing wildlife risks based on food web modeling, they determined that the Coke Point area showed higher risks to wildlife than the Patapsco River background area, and that PCBs and PAHs are the most likely to drive risk in the area (EA, 2011).

The Coke Point risk assessment also addressed potential human health risks at the site. Two risk assessments were performed, with the first considering risks given the

current status of the offshore area, which was "not expected to be frequently used" for recreational purposes (EA, 2011). This assessment analyzed carcinogenic and non-carcinogenic risks to adult, adolescent and child recreational users, and to watermen, for exposure to surface water, sediment, and fish and crab tissue. They concluded that carcinogenic risks exceeded the USEPA screening levels for all groups except child recreational users, but that non-carcinogenic hazards were only in excess of USEPA target levels for child recreational users (EA, 2011). Risk was driven in these cases by dermal exposure to surface water, with contributions from fish and crab tissue consumption – though the analysis noted that carcinogenic consumption levels were comparable to those for the Patapsco River background area, which already has a fishing advisory in place for PCBs (MDE, 2014). Chemicals contributing risks included PAHs and PCBs in fish tissue, and PAHs and dioxins in surface water (EA, 2011). The second human health risk assessment evaluated cumulative risks for future planning at the site, calculating risks based on a theoretical maximum exposure scenario, and concluded that risks exceeded those of the Patapsco River Background Area (EA, 2011).

Results from the Coke Point study, which demonstrated contaminant concentrations leading to elevated ecological and human health risks, brought further attention to nearby Bear Creek, further north from the peninsula but still directly offshore from the Sparrows Point complex. In 2011, Exponent, Inc. (Exponent) produced a draft report for the Chesapeake Bay Foundation (CBF) that concluded the conditions at the Sparrows Point site posed an "imminent and substantive endangerment" to human health and environmental integrity, and that the site warranted further study. "Despite a variety of work having been performed at Sparrows Point…evaluating potential exposure to

contaminants in all of the outlying areas around Sparrows Point is difficult. This is because a systematic characterization of the extent of sediment contamination has not been conducted thus far" (Exponent, 2011). Historical samples had not been collected systematically to represent the entirety of the region, and thus evaluating potential for exposure pathways from direct contact to sediment and from fish and crab consumption was not possible. Given the proximity of Bear Creek to residences, schools, parks, docks and recreational fishing areas, a similar study to the Coke Point analysis is warranted.

In the fall of 2014, EA published their work plan for a Sparrows Point Offshore Investigation, prepared for the Sparrows Point Environmental Trust. Through a combination of field sampling, analytical testing of sediments and porewater, and the calculation of human health and ecological risk based on measured chemical data and modeled surface water concentrations, the study aimed to assess risks at the site. Preliminary results from this study were released in early 2015, indicating ongoing contamination with PAHs, metals and PCBs (EA, 2015). Exponent also continued their assessment of the site, reviewing a database of information on samples from the Sparrows Point area in 2015 to explore their potential for negative effects on human health. Simultaneously, the current study was undertaken to supplement ongoing risk assessments with additional chemistry data, and results from toxicity tests – a measure of environmental quality not included in either of the two recent assessments.

Study objectives

Bear Creek and the larger Baltimore Harbor/Patapsco River system have been subject to numerous environmental studies in the past two decades, but major limitations exist in our understanding of the pattern of toxicity and contamination within Bear Creek. Studies of contamination and toxicity have primarily been performed within the context of larger Harbor-wide studies, with repeated sampling and assessment at a series of –pre-defined locations including the historically polluted Site 28. Both EA and Exponent have proceeded with ecological and human health risk of Bear Creek within the past year. These risk assessments improve upon previous studies with regards to the spatial distribution of sampling sites. However, they focus primarily on the direct offshore area of Sparrows Point, within 100 m of the shoreline. Though these areas are arguably most relevant for human exposure, the areas closest to the shore are comprised of sediments with a higher sand content, which have historically harbored lower contaminant concentrations and have recently been demonstrated to be less toxic than those with lower sand content (Ashley and Baker, 1999; Yonkos, 2015). Our understanding of contamination and toxicity across the width of Bear Creek, and particularly in closer proximity to neighborhoods across the channel from the Sparrows Point complex, is limited. To better assess human health and ecological risks in the area, and to guide management/remedial decisions in the future, a finer-scale assessment of toxicity within the creek is necessary. In addition, we lack knowledge of toxicity and contamination in subsurface sediments at the site. While these sediments are not currently part of complete exposure pathways for ecological or human health risks, the potential for disturbance of the sediments during redevelopment exists. Understanding the characteristics of the

sediments below surface will provide valuable information to guide risk assessment under development scenarios, and to guide future management decisions at the site.

The current risk assessments of Bear Creek are also based upon a comparison of measured contaminant levels in environmental samples with relevant screening levels for specific contaminants in sediment and water. As the history of research in Bear Creek demonstrates, the sediments in the creek contain a complex mixture of numerous potentially toxic agents. Though toxicity to benthic organisms is established in Bear Creek sediments, toxicity identification and evaluation experiments have been inconclusive as to which compounds drive toxicity (Klosterhaus & Baker, 2006). In a complex contaminant mixture with potential additive effects of multiple contaminants, predicting ecological or human health risks based on chemical data may be insufficient. Sediment toxicity assays provide a valuable metric for assessing ecological risks, providing "direct, quantifiable evidence of biological consequences" that can normally only be inferred from chemical analyses (McGee et al., 1999). In the case of Bear Creek, toxicity assays can help to delineate the extent of the impacted area, even though causative agents of toxicity are unknown.

The first objective of this study is to develop a more complete understanding of the spatial extent of toxicity within the Bear Creek system, including examining toxicity depth profiles within sediment cores. Previous studies have focused on sediment contamination near shore rather than in the area across the channel, which is known as an area of recreational use. Two hypotheses related to spatial distribution of toxicity are put forth for this study: first, that toxicity of sediments will decrease with distance from the shore of the Sparrows Point complex, and second, that toxicity will increase with depth

within sediment cores. This depth-related hypothesis is based on our current understanding of Bear Creek as a site with historic contaminant inputs that have decreased in recent years. Based on Mason et al.'s (2004) assessment of sedimentation rates, sediments within the first 80 cm – the depths sampled in this study – are all likely from a time period of intensive human use. Surface samples were collected in an east-west and north-south grid within the Bear Creek channel, and a subset of cores were collected to explore toxicity at depth.

A second objective is to explore the relationship between toxicity and chemical contaminants within this expanded database of samples. To pursue this objective, two novel analytical techniques are explored for contaminant evaluation – an antibody-based PAH biosensor and a handheld X-ray fluorescence (XRF) device. Both these techniques are substantially faster and lower-cost, and require significantly less sample volume and preparation than traditional methods for contaminant analysis. These methods will be discussed in detail in Chapter 2 (PAH biosensor) and Chapter 3 (Handheld XRF), where we explore the correlations between the biosensor and GC-MS and between handheld XRF and ICP-MS as a means to validate these techniques for use at the Bear Creek study site.

By examining toxicity and potentially related contaminant classes in sediments from areas of Bear Creek that have not been previously sampled, this study expands our spatial understanding of the extent and depth of toxicity and contamination at the site. By including toxicity assays, the study supplements ongoing data collection at Bear Creek with additional evidence of direct biological effects from contaminants, providing valuable evidence for future risk assessments of the site. By exploring the use of

handheld XRF and an antibody-based PAH biosensor, we gather further chemical evidence to investigate chemistry-toxicity correlations at Bear Creek specifically, while demonstrating the value of fast and low-cost chemical techniques for pollution mapping at contaminated sites.

MATERIALS AND METHODS

Site description and sample collection

The site assessed in this study comprises the offshore area directly to the west of Sparrows Point. Sampling sites for toxicity tests were identified in a grid spanning the width of the channel and in a rough north to south gradient. (Fig. 1.4). Sites were identified in areas of previously established and/or suspected toxicity, and at increasing distances from the shoreline. Sites were specifically selected in depositional areas with low-sand-content sediments, as these have demonstrated greater toxicity to benthic organisms in the recent past (Yonkos, 2015)

Sediment samples from twenty stations in the region were collected on June 3-4, 2015, according to standard American Society of the International Association for Testing and Materials (ASTM) and USEPA protocols (ASTM, 1994; USEPA, 1995). On June 11, two more sites were sampled (A2 and A'1, Figure 1.2). Sediment samples were collected using a full-size Ponar grab sampler and boat-mounted davit.

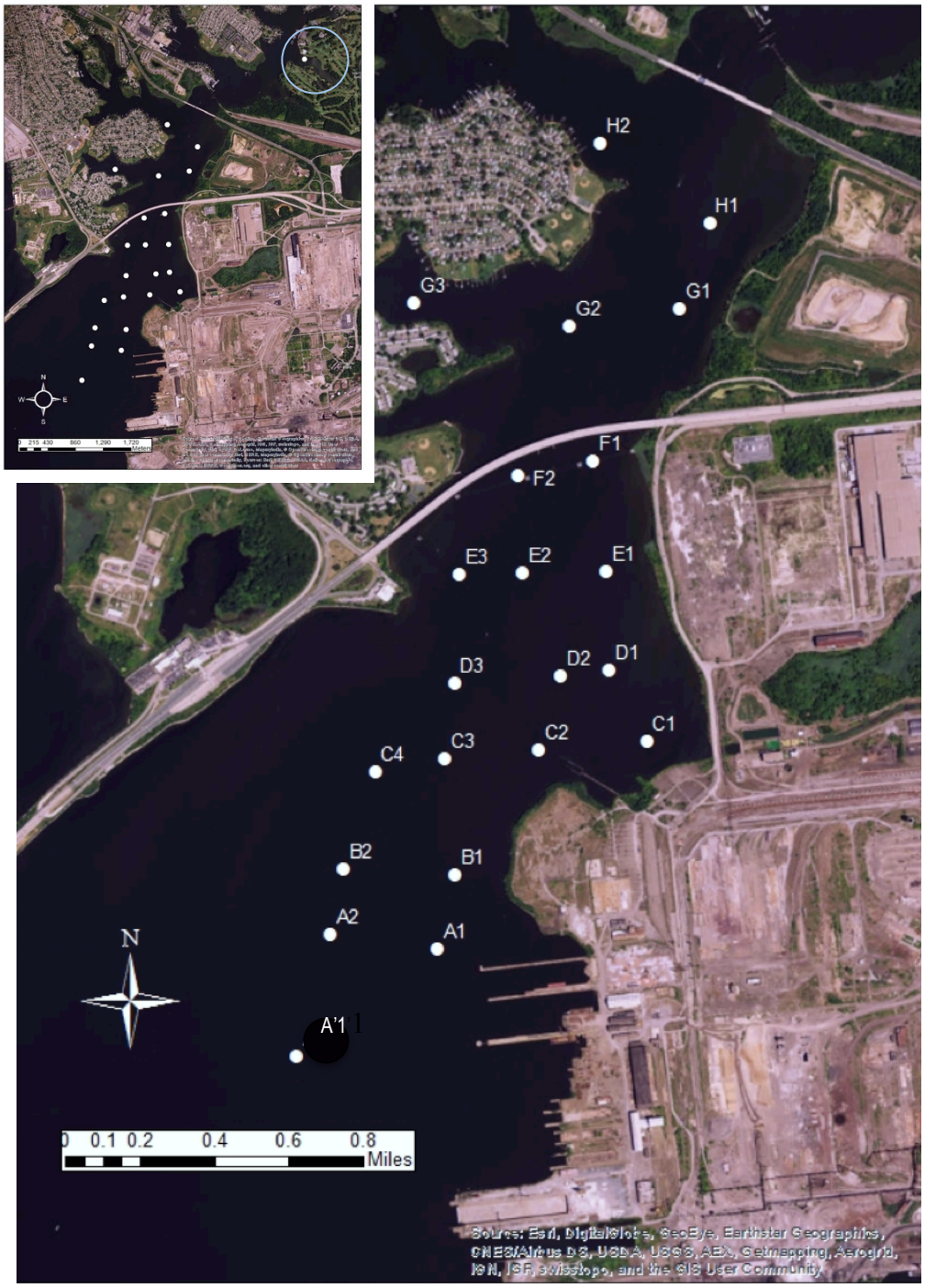


Figure 1.4. Sampling sites in Bear Creek, labeled as transects H through A' Inset, upper left: site circled in blue served as an in-system control sample

Once ponar grabs were brought on board, they were placed in stainless steel metal trays (cleaned with site rinses and acetone between sites). The top two centimeters of multiple Ponar grabs were homogenized to generate testing material, by mixing until homogeneous in texture and color. Sufficient material was collected for both toxicity tests and chemical analyses of sediments.

Sediment samples from twenty stations in the region were collected on June 3-4, 2015, according to standard American ASTM and USEPA protocols (ASTM, 1994; USEPA, 1995). On June 11, two more sites were sampled (A2 and A'1, Figure 1.2). Sediment samples were collected using a full-size Ponar grab sampler and boat-mounted davit. Once ponar grabs were brought on board, they were placed in stainless steel metal trays (cleaned with site rinses and acetone between sites). The top two centimeters of multiple Ponar grabs were homogenized to generate testing material, by mixing until homogeneous in texture and color. Sufficient material was collected for both toxicity tests and chemical analyses of sediments.

Figure 1.5. Collection of surface sediment with Ponar and boat-mounted davit by L. Yonkos and S. Hartzell

To assess spatial variability within the system, four samples were assessed in both homogenized and replicate form. Homogenized samples were the product of mixing several Ponar grabs and then splitting the total sediment sample into five treatment chambers, while replicate samples have five replicates that each represent a single Ponar grab.

Sub-samples for toxicity testing were placed in pre-cleaned 2.5L HDPE containers and held on ice while in the field, then refrigerated at 4^{o}C until processing for sediment tests. Subsamples for chemical analysis were stored in 250-mL certified amber jars and pre-cleaned 1L Mason jars. Reference sediment was collected from Bigwood Cove, a Wye River tributary. An in-site reference sample was also collected from a site within the Bear Creek system but at a historically non-toxic location (Figure 1.5 inset). On September 17th, 2015, repeat samples were obtained from sites C1, D1 and D2 for a second toxicity test.

Core sample collection and processing

Core samples were collected from a subset of six sites (C1, C2, C3, C4, F1 and G1) on July 28th, 2015 and from three further sites (F2, G2, G3) on September 17th, 2015. Core sites were selected based on results from surface toxicity tests. Sediment cores were collected using an in-house fabricated coring device, which was operated using a boat-mounted davit. The coring device was comprised of a 4-inch diameter PVC tube connected to adjustable lengths of 2-inch PVC pipe, used to plunge the core through sediment layers. Inside the PVC system, adjustable lengths of steel rods were attached to a rubber valve to create suction within the column, which held collected samples in place.

To collect cores, the apparatus was attached to the boat-mounted davit and the steel rods were pushed through the PVC system so that the rubber valve was flush with the bottom of the four-inch PVC core casing. The rods were clamped in place, and the tube was lowered into the water. The apparatus was positioned perpendicular to the sediment surface, with the valve resting on the sediment surface. The PVC system was then plunged by hand into the sediment while the valve was held in place at the surface of the sediment. Using the assistance of the boat-mounted davit, the cores were lifted from the water, with the sediment core held intact within the PVC casing.

Figure 1.6. Sediment coring process. Attachment of coring device to boat-mounted davit (left) and use of davit to lower coring device into water (right).

(a)

(b)

Figure 1.7. Two intact sediment cores from Baltimore Harbor, with oily sheen visible on the surface. Cores were occasionally homogeneous in color and texture (a) but many showed clear stratification of layers, showing that sediments were likely not mixed during the plunging and removal process (b) and reflecting potential differences in redox conditions, contaminant profiles, or other sediment characteristics.

Cores were capped for transport to the University of Maryland, and then were stored vertically in a 4^{o}C refrigerator until processing. Before the initiation of sediment tests, cores were removed from the refrigerator and split open length-wise using electric

sheet-metal shears. Cores were segmented by depth, with 10 cm increments from 0 cm to 40 cm and 20 cm increments between 40 and 80 cm. Subsections of cores were stored in Ziploc® bags in a refrigerator at 4^{o}C.

Repeat cores from sites C1, C3 and G1 were collected for GC-MS analysis of sediment porewater on January 9th, 2016.

Surface sediment toxicity tests

Toxicity of sediment samples was investigated using methods described in the *Standard Guide for Conducting 10-day Static Sediment Toxicity Tests with Marine and Estuarine Amphipods* (ASTM, 1992). Test organisms were benthic estuarine amphipod *Leptocheirus plumulosus,* purchased from Chesapeake Cultures, Hayes, VA. Organisms were shipped priority overnight. Organisms were between 2 and 4 mm in length – passed through a 750 μm screen, but retained on a 500 μm screen. Two rounds of toxicity bioassays were performed for surface sediments. The first 10-day exposure began June 9, 2015, with the second starting June 23rd, 2015. Additional surface sediment tests were conducted for sites C1, D1 and D2 (repeat samples) in November of 2015, in order to assess variability of sediment toxicity test results between different sampling events.

Surface sediments, homogenized in 2.5 L buckets, and reference sediment from the Wye River, were processed prior to testing by sieving through a 500 um mesh screen to remove debris, resident amphipods, competitors, and predators. For each sample site, five replicate 1L glass beakers were loaded with 175 mL aliquots of sieved sediment using a stainless steel spoon. Dechlorinated municipal water was aged, aerated and adjusted to 15 ppt salinity to mimic an estuarine environment, and was added to test

chambers for a final volume of 1 L. Baffles were used to minimize sediment suspension during addition of water. Test chambers were arranged in randomized fashion on the testing table (Fig. 1.9).

Prior to test initiation, overlying water was siphoned from chambers and renewed, approximately 24 hours after it had been added. Test chambers were gently aerated using 1 mL plastic pipettes, at approximately 1 bubble/second. Tests were conducted at 23+/-1^{o}C under constant fluorescent lighting. On Day 0, prior to introducing test organisms, general water chemistry (DO, pH and temperature) was measured on all replicates, as well as on Day 9 prior to test conclusion. Ammonia and salinity were also measured on Day 0 and Day 9 on one replicate each. On all other days water chemistry analysis (DO, pH, temperature, ammonia and salinity) was performed on only one replicate per sediment site. Porewater ammonia was also analyzed prior to the beginning of the test. Bulk sediments were placed in 50 mL conical-bottom HDPE centrifuge tubes and were centrifuged at 3,500 rpm for 15 minutes. Separated porewater was decanted and ammonia was analyzed immediately using a LaMotte Smart3 Colorimeter.

Toxicity tests were initiated by loading 20 *L. plumulosus* into each test chamber on Day 0. Organisms were unfed for the duration of the test. Observations of test chambers were made daily. On Day 10 of the test, overlying water and sediments were rinsed through a 500 μm sieve in order to collet and count organisms. Observations were made on numbers of living organisms, as well as on numbers of identifiable bodies.

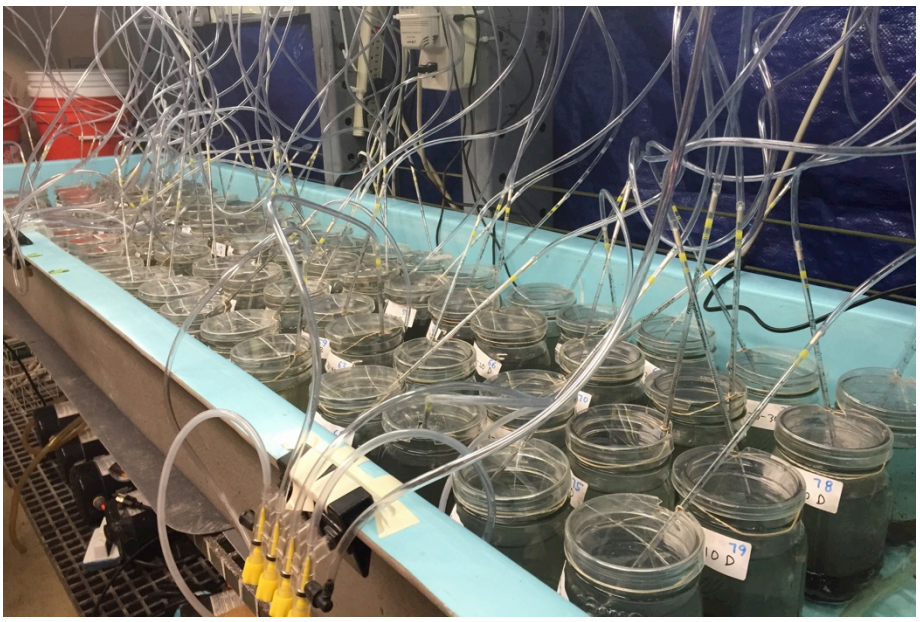


Figure 1.8. Toxicity test table with aerated test chambers.

Core sediment toxicity tests

Toxicity tests for sediment core samples were conducted under the same general parameters as surface sample tests, described above. Given space limitations of the testing table, cores were tested in groups of three. The test for cores C1, G1 and F1 was initiated on August 4th, 2015, while cores C2, C3 and C4 were tested beginning on August 28th, 2015. Deviations from the sediment testing protocol were made due to limitations of sediment availability. While controls were still run in replicates of five, core material was only sufficient to run four replicates at volumes of 150 mL of sediments (control sediments were also run in 150-mL aliquots). Limited material precluded the measurement of porewater ammonia before initiation of test. Water chemistry was conducted as described above. Tests were initiated by introducing 20 *L. plumulosus* into each test chamber on Day 0. A toxicity test under the same parameters

was run on 3 repeat cores from sites C1, C3 and G1 in January of 2016, in order to assess variability of toxicity test results between two separate core sampling events.

Sediment chemical analysis

Subsamples were retained from each surface sediment sample and core segment for analysis of select heavy metals and organic constituents. Metals analysis was performed at the University of Maryland Chesapeake Biological Laboratory, Solomons, MD. Samples were hand-delivered to Dr. Andrew Heyes after being kept on ice. Inductively coupled plasma mass spectrometry (ICP-MS) was used to analyze concentrations of Cr, Zn, Cu, Ni, Pb, V and As in each sample. Samples were weighed and placed in a VWR Scientific Forced Air Oven at 60^{o}C overnight, and were then reweighed to determine dry weight/wet weight ratios. Another subsample of each sediment was placed in an acid-cleaned quartz flask for microwave digestion, using EPA Method 3052. The Milestone EOTHO-EZ uses quartz reaction vessels placed inside PTFD flasks, which are sealed during the digestion. For digestion, 1-2 g of sediment were placed in the vessel with 9 mL of concentrated ultra pure HNO_3 and 2 ml of concentrated ultrapure HCL. The vessel was capped with a loose fitting quartz cap, and placed in the Teflon flask; 5 ml of 30% H_2O_2 was added to the Teflon sleeve and the sleeve was sealed. The sample was heated to 180^{o}C in the microwave for 15 minutes then allowed to reflux for 20 minutes. The samples were then cooled and filtered through Whatman No. 41 filter paper by suction filtration and diluted to required volume with deionized water. These extracts were analyzed using a Hewlett-Packard 4500 Inductively Coupled Plasma-Mass Spectrometer (EPA Method 6020). Standard reference materials and Digestion acid

blanks were analyzed with each batch - NIST 1646a and NIST 1944 (NY/NJ), and standard additions were done to test for interferences (Heyes, personal communication, 2016).

Analysis of ten whole sediment sub-samples was performed for total petroleum hydrocarbons (TPH), polycyclic aromatic hydrocarbons, and polychlorinated biphenyls at Texas A&M University. Samples H1, G1, G3, F1, E1, E2, D1, D2, C1, and C3) were mailed to Dr. Terry L. Wade at Texas A&M using appropriate handling and chain of custody procedures. The sediments were extracted for total PAH, PCB and TPH analysis using methods previously described by Kirman et al (2016). For TPH and PAH measurement, a flame ionization detector was used for quantitative detection. For PCBs, GC-MS in selected ion mode (SIM) was used (Wade, personal communication, 2016).

PAH porewater analysis

Polycyclic aromatic hydrocarbons were also analyzed in porewater of each surface and core sample using an antibody-based biosensor, developed by Drs. Kaattari and Unger at the Virginia Institute of Marine Science in Gloucester Point, VA. To obtain sediment porewater, bulk sediments were centrifuged at the Wye Research and Education Center in 50 mL conical-bottom HDPE tubes at 3,500 rpm for 15 minutes. Porewater was decanted and frozen until analysis. The day prior to biosensor analysis, polymethylmethacrylate beads (Sapidyne) were coated with the antigen, a pyrene-butyric acid-bovine serum albumin conjugate (PBA-BSA). 200 µg of a 1.68 mg/mL stock was added to 200 mg of beads. The antibody used for experiment was 2G8, an anti-pyrene-butyric acid monoclonal antibody produced in mice, sensitive to all three- to five-ring

PAHs (Li et al., 2016). The antibody was frozen down and then fluorescently tagged with AlexaFluor 647. The antibody solution was made with 30 mL diluent to 240 µL mAb.

Antibody solution and antigen-coated beads were loaded into the instrument, and voltages were stabilized using water samples. Spier et al. (2011) describe the automated sample-handling program of the KinExA Inline Sensor. Upon sample introduction, a coated bead pack (approximately 400 µL) was loaded into the flow cell. Next, 400 µL of fluorescently-tagged antibody was loaded, along with 400 µL of water, standard or sample. After mixing within a mixing syringe, half the solution was discarded while the rest was introduced to the bead back loaded into the flow cell, and fluorescence was recorded by the instrument. As Spier et al. (2011) write, the fluorescence signal "was based on competitive exclusion by free PAHs in the sample such that the amount of [antibody] bound to the antigen-coated beads within the flow cell was inversely proportional to the concentration of three- to five-ring PAHs in the sample." Fifty percent dimethylsulfoxide in DI water was used for automated rinsing between samples. Samples were loaded 7 at a time.

Standards, prepared with phenanthrene in methanol added to DI water at concentrations of 0.5 µg/L, 1 µg/L, 1.5 µg/L, 2.0 µg/L and 2.5 µg/L, were used to create a calibration curve to determine porewater concentrations. After stabilization with water and analysis of standards, samples were introduced. Prior to introduction, samples were filtered through a 0.45 µm filter (Millipore). To bring samples within the calibration range of the instrument, many samples were diluted from between a 1:1 dilution and a 100:1 dilution with deionized water.

Geostatistical analysis

Kriging was performed using surface sample data as the input layer, in order to predict survival and contaminant concentrations throughout the system. Multiple combinations of kriging method and semivariogram models were compared, and Ordinary Kriging using a Stable semivariogram model was chosen for mortality data and applied to each contaminant class. Probability kriging surfaces were created using ArcGIS's Empirical Bayesian Kriging tool to predict likelihood of exceeding predicted effects concentrations for sediments.

Using ArcGIS Geostatistical Analysis tools, spatial interpolation (kriging) was performed for mortality data and for concentrations of chemicals within sediments. As a first step in the kriging process, the Geostatistical Wizard was used to examine several semivariogram models, which, when fit to a set of data, measure the strength of a statistical correlation as a function of distance (ESRI, 2011). Several semivariogram models were compared, and the prediction error statistics were computed for each. Unbiased prediction errors are indicated by a mean standardized prediction error near zero, and valid predictions are indicated by a root-mean-squared standardized value near 1 (ESRI, 2011). ArcGIS's Empirical Bayesian Kriging (EBK) tool, which automatically optimizes parameters by repeatedly estimating new semiovariogram models, provided the best values (Krivoruchko, 2012). However, Universal Kriging using a Stable semiovariogram was chosen to create prediction surfaces, so that settings could be documented and maintained between factors. To create a mortality prediction surface that covered a wider range, site 34 was included in the analysis.

Statistical analysis

The only endpoint for the *Leptocheirus* test was survival. Data were found to be non-normally distributed, and so statistical analysis of survival data was performed on square-root arcsine transformed data using the Kruskal-Wallis ANOVA on Ranks test with Dunn's post-hoc comparisons between treatment means and means from both the Bigwood Cove control site and the in-system control. Statistical tests were run using SigmaPlot version 12.0 software, with significance set at a minimum probability level of $p = 0.05$. Samples were also analyzed individually within their respective transects. For analysis of core samples, each core was analyzed individually, with segments compared to the respective test's control sample with the lowest survival. These data were also non-normal, necessitating a Kruskal-Wallis test with Dunn's post-hoc comparisons between core segments and treatment means. To assess the relationship between toxicity and chemistry, a Pearson correlation coefficient test was computed for survival data and for each chemical parameter.

RESULTS

Sediment toxicity tests – surface

For each test, average overlying water quality parameters in amphipod toxicity test chambers remained within acceptable limits for the duration of the test. All control sites had average values greater than or equal to 90%, above the 80% required for a valid 10-day acute toxicity test (ASTM, 1992). Surviving organisms were counted from each test chamber, and the mean percent survival was calculated

by averaging the survivals from each test chamber (Tables 1.1 and 1.2). One treatment replicate from site A1 was accidentally discarded during the test takedown, so the mean percent survival was calculated from only four replicates. Mean survival values are presented as percentages out of 100.

Table 1.1. Mean percent survival of *Leptocheirus plumulosus* in 10-day sediment toxicity test. Tests sites were compared to two control samples, one from within the Patapsco River system in an un-impacted area (Control (IS)).

Site	Mean % Survival	Standard Deviation of Replicates
CONTROL	98	2.74
CONTROL (IS)	90	17.32
A1	**34**	12.50
E1	**8**	8.37
E2	**42**	12.04
E3	67	14.83
F1	**28**	13.04
F2	**54**	17.10
G1	69	14.75
G2	67	9.08
G3	91	8.22
H1	76	16.73
H2	83	9.75

Mean percent survivals in the test samples ranged from 8% to 91%, with standard deviations ranging from 2.74 to 17.10. Of the eleven test sites included in the first

toxicity test, five – A1, E1, E2, F1, and F2 - showed significant differences from the Bigwood Cove control sediment. All sediments that differed significantly from the Bigwood Cove sediment were also significantly different from the in-system control sediment, except for sample F2 (54% survival). For this test, sediments with above 60% survival did not show significant differences from either control (See Appendix B).

Table 1.2. Mean percent survival of *Leptocheirus plumulosus* in 10-day sediment toxicity test. Tests sites were compared to two control samples, one from within the Patapsco River system in an un-impacted area (Control (IS)).

Site	Mean % Survival	Standard Deviation of Replicates
CONTROL	99	2.24
CONTROL (IS)	98	2.74
B1	**13**	11.09
B2	49	21.62
C1	**1**	2.24
C2	**9**	8.94
C3	29	17.82
C4	50	14.14
D1	**0**	0.00
D2	**2**	2.74
D3	**5**	5.00
A2	61	13.87
A'1	42	29.07

Mean % survivals in the second set of test sites ranged from 0% to 61%, with standard deviations between 0 and 29.07 – reflecting a wide range of in-site variability.

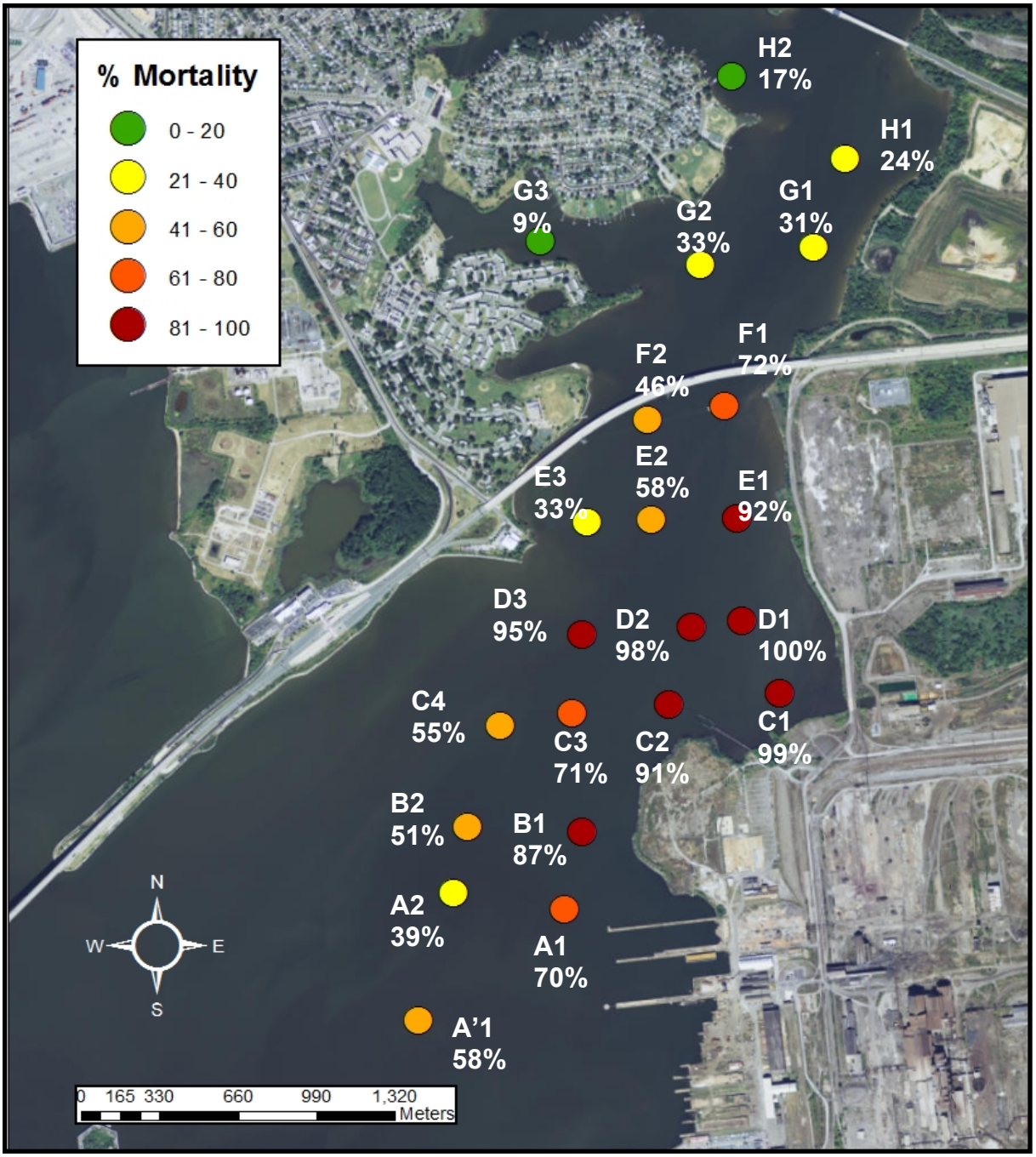


Figure 1.9. Map of *Leptocheirus plumulosus* mortality in 10-day sediment test. % survival represented by color gradation, with 81-100% mortality in dark red.

Of eleven test sites included in the second toxicity test, six sites – B1, C1, C2, D1, D2, and D3 – were significantly different from both control sediments. sediments with 42% survival and greater were not statistically distinguishable from the control sediment (Appendix B).

Samples that caused the highest lethality were generally located near to the shore of the Sparrows Point complex. Near-shore sites from transect A to transect F showed statistically significant mortality, with significant mortality extending further into the channel in transects C, D, E and F. Due to nonparametric statistical analysis, only sites with less than 50% survival were significantly different from control sites. However, a number of sites were observed with mortalities between 20 and 40% (yellow) and between 40 and 50% (light orange).

To assess spatial variability within the system, four samples were assessed in both homogenized and replicate form. Homogenized samples were the product of mixing several ponar grabs and then splitting the total sediment sample into five treatment chambers, while replicate samples have five individual replicates that each form a single ponar grab. For site E1, the coefficients of variance were nearly equal between homogenized and replicate samples (Table 1.3). For site G1, the coefficient of variance for the replicate samples was lower than that for the homogenized samples, indicating that replicate survival may vary even when replicates are drawn from a homogenized batch of sediments. The differences observed between the two tests for Sites C3 and D2 indicate some degree of spatial variability in mean survivals, though replicates were similarly distributed within each treatment.

Table 1.3. Replicates vs. homogenized samples at four sites. Average % mortality and coefficients of variance are reported.

Site	Sample Form	Average % Mortality	Coefficient of Variance
E2	Homogenized	58	21
	Replicate	54	24
G1	Homogenized	31	48
	Replicate	33	30
C3	Homogenized	71	25
	Replicate	53	54
D2	Homogenized	98	3
	Replicate	90	10

To assess temporal and/or spatial variability within the Bear Creek system, several repeat samples were collected and analyzed in a separate toxicity test. (Table 1.4)

Table 1.4. Survival comparison of sediments re-collected from sites C1, D1 and D2

Site	Survival Test 1 (%)	Survival Test 2 (%)	Distance
C1	1	1	2.9 meters
D1	0	5	6.1 meters
D2	**2**	**48**	10.2 meters

While sites C1 and D1 were similar to results from the original sediment toxicity test, site D2 differed greatly. Distances between the original and re-collection sites were measured in Google Earth based on GPS latitude and longitude coordinates, and all distances were in the same range, indicating that substantial variation may exist within relatively small distances in Bear Creek.

Sediment Toxicity Tests– Cores

Averages of control replicate survivals were at or above the 90% survival required for a valid toxicity test. Originally, two control treatments were included in each core test, because 18 treatments (3 cores x 6 segments per core) were run in each test. However, cores were ultimately analyzed separately from one another. For each statistical test, the core was run against the weaker of two controls for that test. Because data was not normally distributed, all sediment cores were analyzed nonparametrically (Appendix B).

Cores within the system showed a variety of vertical patterns. Cores G1, F1, C1 and C2, near the shoreline, showed substantial toxicity (81-100%) throughout the entirety of the column. When compared to the surface grab samples, which are comprised of the top two centimeters of multiple sediment grabs, cores G1 and F1 are substantially more toxic below the surface than they are at the surface. Cores G2 and F2 also show a similar pattern, with a moderately toxic surface layer underlain by a relatively nontoxic top ten centimeters, with increasing toxicity at depth within the core. Cores G3 and C3 show a roughly similar pattern, with toxicity results peaking in the middle section of the core, and improving slightly at depth. Core C4 showed the most reduction of toxicity at depth,

with only the top 20 cm exhibiting substantial toxicity. This distribution of toxicity within cores also reveals horizontal patterns within the channel. From the points near sparrows point to the middle of the channel, toxic responses are substantial for the entire length of the core. As we move across transects, however, toxic responses decrease at depth.

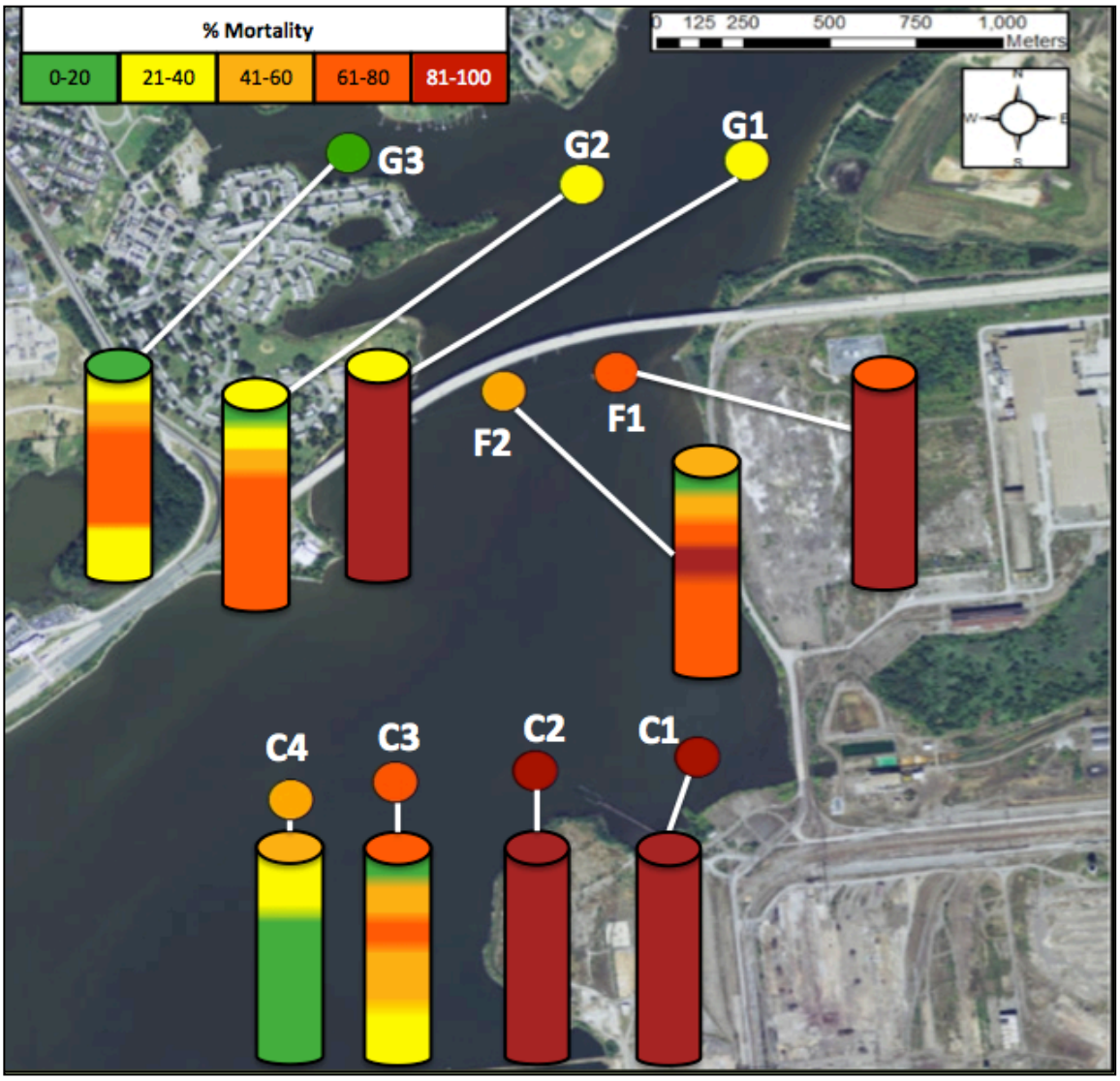


Figure 1.10. Map of % mortality within sediment cores. Circular tops of cores represent toxicity observed in the top 2 cm at the site (homogenized ponar samples)

Table 1.5 Mean Percent survival of test organisms within core segments

Depth	C1	C2	C3	C4	F1	F2	G1	G2	G3
0-10	0	8	85	74	6	86	18	88	63
10-20	0	4	51	78	0	49	3	60	58
20-30	0	3	45	76	0	33	0	53	33
30-40	0	16	25	85	1	13	0	36	25
40-60	0	14	51	89	3	35	1	21	29
60-80	0	6	60	94	0	33	3	21	61

Core C1 caused 100% lethality for the entire depth sampled. In core C2, survivals were uniformly below 20%, but higher values were observed between 30 and 40 cm (Table 1.5). Core C3 showed highest survivals at the surface, with moderate lethality at depth while in Core C4 survivals were generally high with improvement at depth. Cores F1 and G1 were largely toxic, with highest survivals in the top segments. F2 and G2 show a similar pattern, with high survivals in the top segments and decreasing survival at depth, while G3 exhibits decreasing survival with a band of higher survival at 60-80 cm.

Mean percent survivals were also compared between the original cores collected, and repeated cores collected in January of 2016 to obtain sediment volume necessary for GC analysis (Table 1.6). Cores were taken from the same approximate GPS points as the original set, indicating that small-scale spatial variability is present at this study site.

Table 1.6. Mean % Survival compared between two core sampling events at the same sampling sites.

Core Section	Mean % Survival – Test 1	Mean % Survival – Test 2	Distance Between Cores
C1 0-10	0	5	
C1 10-20	0	3	
C1 20-30	0	0	
C1 30-40	0	0	3.01 meters
C1 40-60	0	0	
C1 60-80	0	0	
C3 0-10	85	71	
C3 10-20	51	50	
C3 20-30	45	51	9.6 meters
C3 30-40	25	33	
C3 40-60	51	46	
C3 60-80	60	43	
G1 0-10	18	70	
G1 10-20	3	36	
G1 20-30	0	24	5.23 meters
G1 30-40	0	11	
G1 40-60	1	25	
G1 60-80	3	46	

Sediment chemical analysis

Results are reported for chemical analysis of metals in surface samples by ICP-MS, with concentrations given per dry weight of sediments (Table 1.12). Both core segments and surface samples were also analyzed by handheld XRF, with results presented in Chapter 3. Sediment metal concentrations were compared to consensus-based probable effects concentrations, above which adverse ecological effects would be expected (MacDonald et al., 2000). They were also compared to effects range low (ERL) and effects range median (ERM) values for toxic sediments, which define the concentration below which effects are rarely observed (ERL) and

the concentration above which effects would frequently occur (ERM) (Long and Morgan, 1991).

Table 1.7. Concentrations of select metals in Bear Creek sediments compared to PEC, ERM and ERL values. Concentrations in bold exceeded the consensus-based PEC.

Station	Metal conc. (µg/g)						
	Zn	Cr	Cu	Ni	Pb	As	Cd
H2	**1524.7**	**965.6**	**196.15**	47.6	**230.75**	30.15	**13**
H1	**1362.8**	**913.1**	**164.2**	**58**	**203.3**	29.4	**7.6**
G1	**1513.6**	**987.2**	**155.5**	**65.6**	**193.9**	**33.5**	**12.7**
G2	**1585.7**	**1066.1**	**183.3**	**60.1**	**239.2**	29.5	**7.4**
G3	**2229.8**	**1220.1**	**315.5**	**82.6**	**351.5**	**43.6**	**10.3**
F1	**1455**	**944.1**	145	**51.2**	**178.8**	23.3	**8.3**
F2	**1389.45**	**931.8**	**170.8**	21.65	**239.5**	11.9	2.55
E1	**1569.2**	**1153.3**	**163.8**	**63.1**	**176.9**	28.5	**6.6**
E2	**1168.3**	**853**	147.8	**53.4**	**189.5**	21.6	**5.0**
E3	**920**	**547.4**	79.2	32.6	**184.5**	**35.8**	4.1
D1	**2294.9**	**2685.65**	**228.65**	**118.25**	**224.95**	**50.8**	**9.3**
D2	**1451.4**	**1281.8**	**176.2**	**71.9**	**179.2**	27.8	**5.5**
D3	**1145.55**	**882.1**	138.3	47.85	**184.15**	31.25	**6.45**
C1	**1415.95**	**3196.05**	**212.95**	**127.7**	96.1	25.4	**5.05**
C2	**1655.3**	**1240.3**	**176.6**	**76.1**	**188.8**	**33.6**	**6.9**
C3	**983.3**	**720.3**	139.6	**59.3**	**175.4**	27.3	**5.3**
C4	**1378.7**	**799.8**	**186.7**	43.5	**485**	**62.1**	4.8
B1	**2000.9**	**1720.9**	**343.4**	**69.2**	**383.7**	**44.7**	**13.2**
B2	**986.2**	**856.7**	114	37.1	**259.4**	**55.6**	**2.5**
A1	**705.8**	**601.2**	144.3	**62.4**	**146.5**	30.3	4.9
A2	**604.8**	**467.4**	98.5	38.6	**150.9**	24.7	**2.2**
A'1	**774.8**	**614.85**	142.45	43.85	**189.95**	28.3	**2.55**
PEC	459	111	149	49	128	33	5
ERL	150	81	34	21	47	8.2	1.2
ERM	410	370	270	52	220	70	9.6

Zinc concentrations ranged from 600 µg/g to over 2000 µg/g with an average of 1368.9 µg/g, and exceeded sediment screening guidelines throughout the entire system. Chromium results were similar, with concentrations ranging from 467.4 µg/g to upwards of 3000 µg/g and an average concentration of 1120.4 µg/g . For both these metals, the

lowest concentrations were found in site A2. Copper, nickel, and cadmium all exceeded sediment screening guidelines at over half the sites sampled, while arsenic showed excess levels at eight sites. The average copper concentration was 173.8 µg/g, which exceeds the PEC value. Nickel concentrations averaged to 60.5 µg/g, exceeding all three screening benchmarks. Arsenic values had an average of 33.1, equal to the PEC value. Lead was elevated above sediment screening guidelines at every site but site C1.

Concentrations were obtained for total PAHs in sediment porewater via PAH biosensor analysis. Total PAHs, total PCBs, and TPH in whole sediments were determined in a subset of samples at GERG laboratory, Texas A&M. These concentrations are compared to consensus-based PEC values (Macdonald et al., 2000), as well as to ERL and ERM values (Long and Morgan, 1991). TPH values are compared to the SL-SQS, a proposed screening level sediment quality standard for the protection of benthic organisms (Inouye, 2014). For all contaminant classes, several sites within Bear Creek exceeded established screening levels.

Table 1.8. Concentrations of organic contaminants in Bear Creek sediments. Bold values exceed PEC level; italicized values exceed SL-SQS

Station	PAH Porewater (µg/L)	PAH Total Sediment (µg/g)	PCB Total Sediment (µg/g)	TPH Total Sediment (µg/g)
H2	0.41	ND	ND	ND
H1	0.47	12.3	**0.40**	*5779*
G1	0.51	13.8	**0.50**	*7504*
G2	0.39	ND	ND	ND
G3	0.39	16.3	0.39	*6204*
F1	0.51	15.3	**0.63**	*11745*
F2	0.5	ND	ND	ND
E1	0.7975	16.2	**0.76**	*19254*
E2	0.83	18.4	0.37	*5680*
E3	0.38	ND	ND	ND

D1	2.32	**45.0**	**1.16**	*25794*
D2	0.51	15.7	**0.44**	*22479*
D3	0.3	ND	ND	ND
C1	4.37	**49.9**	**1.09**	*54684*
C2	0.9	ND	ND	ND
C3	0.46	**23.3**	0.38	*5759*
C4	0.55	ND	ND	ND
B1	6.72	ND	ND	ND
B2	0.43	ND	ND	ND
A1	0.62	ND	ND	ND
A2	0.42	ND	ND	ND
A'1	0.64	ND	ND	ND
PEC	-	22.8	0.68	
ERL	-	3.4	0.023	
ERM	-	35.0	0.4	
SL-SQS	-			3600

PAH concentrations in sediment porewater ranged from 0.3 μg/g to a maximum of 6.72 μg/g, with an average of 1.06 μg/g. Only a fraction of the total sediment PAHs were reflected in the porewater; total sediment concentrations of PAHs ranged from 12.3 μg/g to 49.9 μg/g, with an average of 22.62 μg/g. For total PAHs, sediment guidelines were exceeded in three out of ten sites. PCBs showed values ranging from 0.37 to 1.16, with an average of 0.612. Sediment quality guidelines were exceeded in seven out of ten sites. For total petroleum hydrocarbons, all sites exceeded the SL-SQS guideline, with an average of 16,488.2 μg/g.

Toxicity and relationship to contaminant classes

Correlations between specific contaminant classes and observed mortality were explored through a Pearson Product Moment Correlation, and relationships observed are summarized in Table 1.18. Mortality showed a significant positive correlation with chromium, nickel, PAHs in porewater and total sediment, PCBs, and TPH. Though zinc,

copper, and arsenic all had positive R values, the correlations were not statistically significant.

Table 1.9. Pearson Product Moment Correlation results for mortality and its relationship to various classes of contaminants. R values and p values are presented. $[PAH]_{pw}$ represents the concentration of total PAHs in porewater, while $[PAH]_{sed}$ represents the concentration of total PAHs in whole sediment. Entries in bold were statistically significant.

		Mortality	$[PAH]_{pw}$	[Cr]	[Zn]	[Cu]	[Pb]	[Ni]
$[PAH]_{pw}$	**R**	**0.45499**						
	p-value	**0.03337**						
[Cr]	**R**	**0.58566**	**0.68521**					
	p-value	**0.00419**	**0.00043**					
[Zn]	**R**	0.16223	0.39992	**0.60881**				
	p-value	0.47073	0.06516	**0.00264**				
[Cu]	**R**	0.10211	**0.65255**	**0.55035**	**0.81613**			
	p-value	0.65114	**0.001**	**0.00796**	**3.64E-6**			
[Pb]	**R**	-0.26155	0.17969	-0.07061	**0.43017**	**0.5472**		
	p-value	0.23968	0.4236	0.75486	**0.04568**	**0.0084**		
[Ni]	**R**	**0.53626**	**0.52603**	**0.88721**	**0.57666**	**0.53053**	-0.15941	
	p-value	**0.01009**	**0.01192**	**3.75E-8**	**0.00496**	**0.01108**	0.47857	
[As]	**R**	0.02283	0.17981	0.15097	0.32188	0.30859	**0.69676**	0.16527
	p-value	0.91968	0.42329	0.50244	0.14407	0.16232	**0.00031**	0.46234

Table 1.10. Pearson Product Moment Correlation results for mortality and its relationship to various classes of contaminants. R values and p values are presented. $[PAH]_{pw}$ represents the concentration of total PAHs in porewater, while $[PAH]_{sed}$ represents the concentration of total PAHs in whole sediment. Entries in bold were statistically significant.

		Mortality	TotalPCB	TPH
TotalPCB	**R**	**0.71801**		
	p-value	**0.01936**		
TPH	**R**	**0.72567**	**0.80442**	
	p-value	**0.01752**	**0.00502**	
TotalPAH	**R**	**0.64858**	**0.85518**	**0.81288**
	p-value	**0.0425**	**0.00161**	**0.00425**

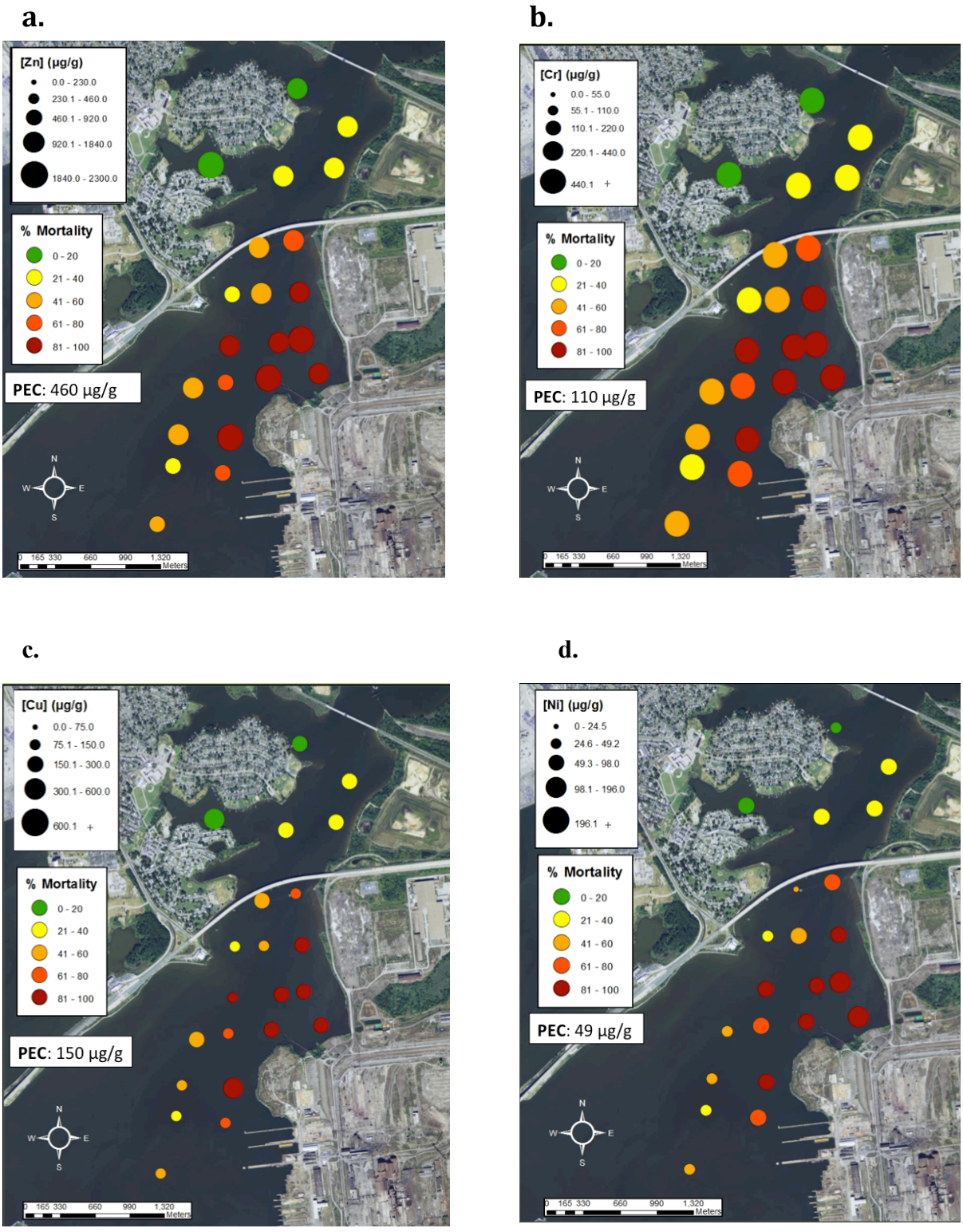


Figure 1.11. Measured concentrations of zinc (A), chromium (B), copper (C), nickel at the stations tested for sediment toxicity. Marker size represents concentration of each metal, while marker color represents observed toxicity at the site.

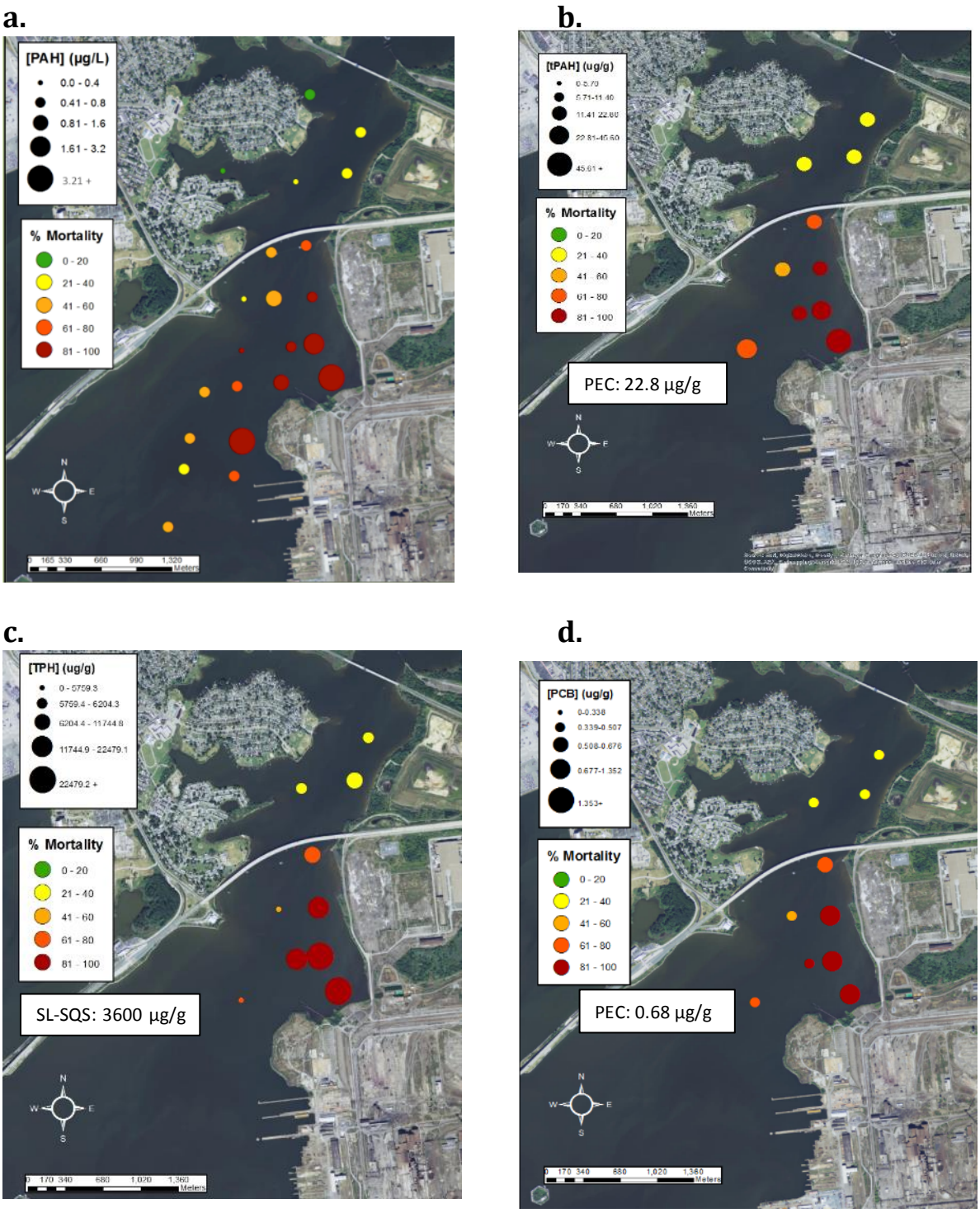


Figure 1.12. Measured concentrations of total PAHs in sediment porewater (A), and total PAHs (B), total petroleum hydrocarbons - TPH (C), and total PCBs (D) in a subset of ten sediment samples. Marker size represents concentration of each metal, while marker color represents observed toxicity at the site.

For both inorganic and organic contaminants, sediments in Bear Creek frequently exceeded the reference values for sediment quality, indicating the potential of this contaminants to cause toxicity within the system. Positive correlations were observed between multiple contaminant classes and mortality, and contaminants also tended to covary with one another.

Geostatistical Analysis

Using ArcGIS Geostatistical Analysis tools, spatial interpolation (kriging) was performed for mortality data and for concentrations of chemicals within sediments. Universal Kriging using a Stable semiovariogram was chosen to create prediction surfaces, so that settings could be documented and maintained between factors. To create a mortality prediction surface that covered a wider range, site 34 was included in the analysis. Results of kriging performed for survival data are presented in Figure 1.13, while kriging results for metals data are presented in Appendix C. For survival data, the prediction surface reveals a hub of toxicity at the outlet location of Tin Mill Canal, at the shoreline within the C and D transects. The prediction surface indicates that the northern boundary of sediment contamination may be consistent with the H transect, while the boundaries to the south and the west have not yet been defined.

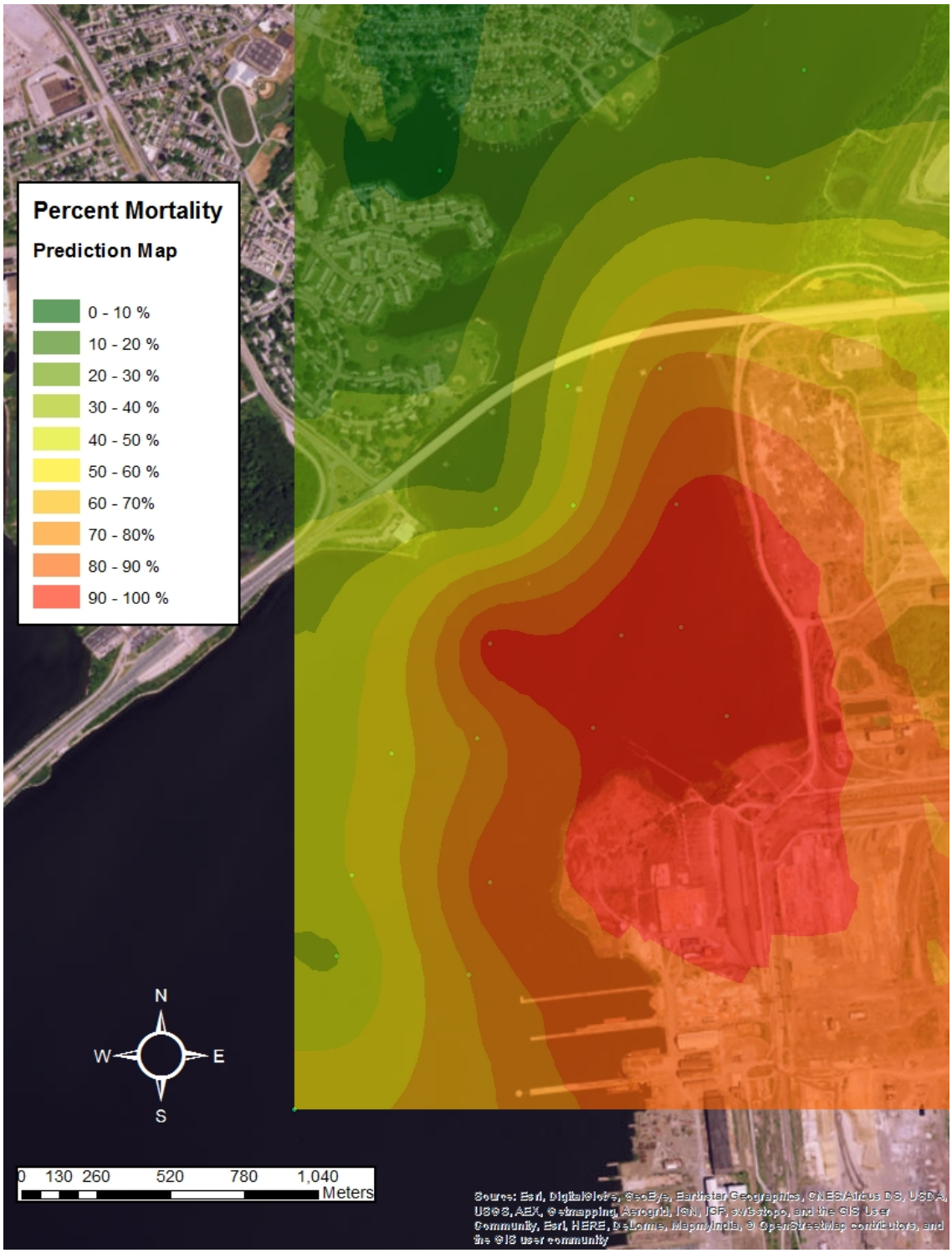


Figure 1.13. Prediction surface for percent survival within the Bear Creek system.

DISCUSSION

Results from the spatial assessment of surface toxicity largely confirmed the first hypothesis of the study; that sediment toxicity would decrease with distance from the shore of the Sparrows Point industrial complex. Toxicity generally decreased in each transect, though in Transect D, toxicity was near 100% across the entire width of the channel. In this region, it is apparent that sediments are highly toxic throughout the entire depositional region of the channel; samples were exclusively collected for testing in these depositional areas, and not in sandier portions of the channel.

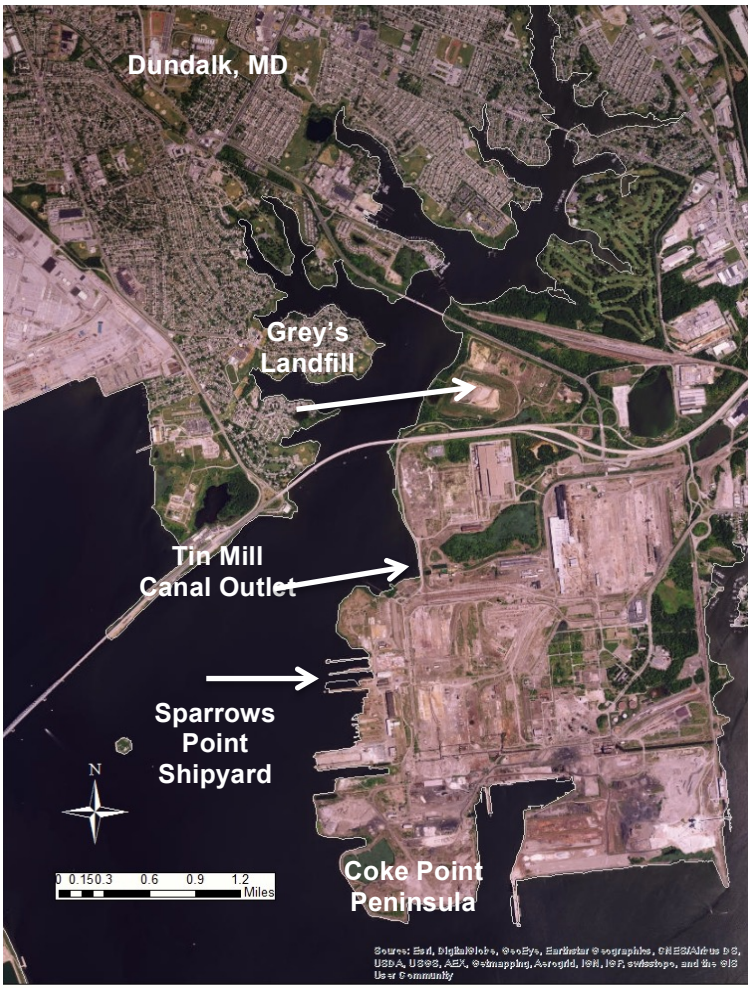


Figure 1.14. Potential sources of contaminants at Sparrows Point

To the north and south of the D transect, the apparent hub of toxicity in the system, mortality decreases. These observed spatial patterns of toxicity within Bear Creek can be related to landforms in the on-site area. Most particularly, the hub of toxicity demonstrated in the C and D transects is likely a result of the outlet of Tin Mill Canal, a constructed swale that drains about 800 acres from the Sparrows Point site (EnviroAnalytics Group, 2015). This canal likely serves as an ongoing source of contaminants, both organic and inorganic, to the system. From this hub of toxicity and contamination, toxic effects in benthic organisms tended to decrease to the north and south within Bear Creek. The Sparrows Point Shipyard, located south of Tin Mill Canal, may also be a source of historic contamination from a history of waste disposal related to shipping activities (Chesapeake Bay Foundation, 2009). To the north of Tin Mill canal is the unlined Grey's Landfill, another source of organics and heavy metals to the system. (Chesapeake Bay Foundation, 2009). In general, toxicity results improved to the north within Bear Creek, with minimal toxicity noted upstream of these major on-site sources. While the creek is tidal, a net flow is still expected towards the main channel of the Patapsco, so contaminant transport is more likely to happen from north to south.

Statistically, sites C1, C2, D1, D2, D3, E1, E2, F1, F2, B1, and A1 were significantly different from the control sediments analyzed. Since nonparametric analyses are less powerful than their parametric equivalent, this is likely a conservative estimate of toxicity. Samples were also analyzed statistically in their specific transects. Statistics aside, trends are apparent within the Bear Creek system, with sediment mortality decreasing with distance from the shore. The results of this study agree with previous

analyses of toxicity within Bear Creek, which showed near-total mortality in the direct offshore area of Sparrows Point (McGee et al., 1999).

An assessment of nine sediment cores to depths of 80 cm provided a depth profile for toxicity at the site. Similar statistical challenges arose with assessing core survival data, because all cores except C3 had to be analyzed nonparametrically (see Appendix B). A relative examination of mortality between core segments is more useful in this case than statistical analysis. In general, toxicity in surface grabs (the top 2 cm of sediment) was not characteristic of toxicity at depth. For most cores, toxicity at the surface was substantially lower than that observed in deeper segments of the core. However, toxicity did not uniformly increase with depth as the study's hypothesis predicted. As seen in Figure 1.8, each core showed a unique distribution of toxicity at depth. Core C1, C2, F1 and G1 were toxic throughout the depth of the core, though in cores F1 and G1 the surface (top 2 cm) were distinct from the sediment beneath. G1 and F1 showed higher survival within the top 2 cm than that revealed at depth. Cores G2, G3 and F2 showed more moderate toxicity at depth, but they continued a pattern observed in G1 and F1. In these cores, the top 2 cm were not characteristic of toxicity at depth, where mortality increased.

Most notable among these cores is Core G3, where the surface segment had less than 20% mortality but sediments at depth showed up to 80% mortality. Core C4 was unique, showing survival increasing at depth. Throughout this core, survival was high, indicating that C4 may mark the western boundary of subsurface contamination. The possibility of other depth-related factors contributing to toxicity within cores cannot be ruled out. During toxicity testing, ammonia levels in overlying water were monitored,

and tended to increase with depth within each core. According to the USEPA (2001), overlying water ammonia for benthic organisms should not exceed 60 mg/L at test initiation, and should not exceed 16 mg/L at the end of the test. For surface tests, ammonia concentrations were low. However, in core tests several treatments, most notably in core C1, exceeded 16 mg/L throughout the test and at test completion. It is possible that ammonia toxicity contributed to the increased mortality of organisms in deeper segments. However, the decreases in survival at depth were not uniform throughout the system, and Core C4's high survival seems to validate the approach of applying toxicity tests at depth.

The initial hypothesis put forth in this study - that toxicity would uniformly increase with depth - was not supported by the data presented. In general, the data support a conclusion that toxicity at the surface cannot be used adequately to predict toxicity at depth. Though each core was unique, within the system more toxic sediments were generally overlain by less toxic layers. This could be explained by several potential mechanisms. It is possible that the less toxic sediments on the surface reflect decreased pollutant inputs over time within the system. However, decreases in toxicity at various points within the core may also be reflective of changes in the sedimentation rate at the site. Additionally, sediment resuspension rates may vary from place to place within the system, with shallow areas more prone to resuspension than deep areas. Dating of cores could yield more information about the age of sediments and thus the sedimentation rate at the site. Additionally, lateral movement of sediments within the system may complicate the understanding of each core as a chronological account of inputs at that site.

Mason et al. (2004) explored the sedimentation rate and age of sediments at Site 28, illuminating the complexity of this site. They report generally increasing sedimentation rates at the site, with a major peak around 1940, another peak in 1951, and a drop in sedimentation after that point; they relate these peaks to potential filling activities at the Sparrows Point shoreline by industry (Mason et al., 2004). They also note that sedimentation rates are high due to the depositional nature of the estuarine environment in Baltimore Harbor. In all, the 2m core they collected for the study represented about 100 years (Mason et al., 2004). It is reasonable to assume that the entire length of our cores in the near-shore area represented a time where Bear Creek was impacted by industry. However, a more thorough exploration of core ages within the system would be illuminating, especially since our toxicity and contamination results show a potential depth limit to contamination at the western edge of the sampling region. Sedimentation rates, though high in the near shore area, may be quite variable throughout the rest of the channel.

Three experiments within this study assessed the potential spatial variability of toxicity within the Bear Creek system. Toxicity tests were run on sediments from three sites in two different manners – multiple grabs homogenized, and multiple grabs stored separately with each run as a replicate. Coefficients of variance were similar between homogenized samples and replicate samples, indicating that homogenization of multiple grabs does not mask spatial variability in the site at the level of adjacent collection. However, some evidence of spatial variability emerged in the homogenization-replicate comparison, where site C3 showed different mean survivals (29% in homogenized vs. 47% in replicates). A second assessment of spatial variability analyzed a second set of

collected samples for three sites, where samples were gathered within 10 m of the previous sampling location. Sites C1 and D1 showed good agreement between the first test and the repeat test, while D2 showed profound differences, with a second-test survival of 48% compared with an initial 2%. This speaks to the potential for substantial variability of toxicity either over short time periods or across seasons (unlikely within these buried sediments) or small distances (likely). A similar test was conducted with core segments, recollected at sites C1, C3 and G1. Cores C1 and C3 showed relatively consistent results between the two tests. The most profound differences are seen in Core G1, where the top section in the first test produced substantially more lethality than in the second test. Survival was seen in all core segments of the second test, while in the first survival dropped to 0% between 20 and 40 cm. These tests together indicate the potential fairly small-scale spatial variability within the study site.

This study also assessed concentrations of several key metals within surface sediments and core segments. Metals within surface sediments exceeded predicted effects concentrations at each site. Zinc and chromium uniformly exceeded predicted effects concentrations throughout the system, and lead exceeded the PEC in all but one site. Copper, cadmium and nickel were all in excess of the PEC values in more than 50% of sites tested. In Table 1.14, PAH concentrations in porewater of surface sediments are displayed, as determined by a PAH biosensor. These values were uniformly low as compared with total sediment PAHs, indicating that only a portion of total PAHs are available in the sediment porewater. Total PAHs, along with PCBs and TPH, showed exceedances within the system. A Pearson Product Moment Correlation was performed to assess relationships between mortality and contaminant concentrations. Of the

constituents assessed, mortality showed a significant positive correlation with nickel, chromium, PCBs, and TPH, though it is difficult to say whether any one of these constituents could drive toxicity.

The Pearson Product Moment Correlation not only highlights potential relationships between contaminants and toxicity, but also reveals that contaminating constituents tend to covary within the system. No single constituent can be singled out as a likely driver of toxicity, as many contaminants vary along with one another. Establishing cause and effect relationships between contaminant classes and toxicity is difficult in complex environmental samples where multiple contaminants may be to blame. Additionally, unmeasured contaminants at the site may be primary drivers of toxicity (Klosterhaus & Baker, 2006).

Though results from this investigation agree with McGee et al., (1999) in identifying heavy metals as a potential driver of toxicity, a thorough assessment of metals toxicity cannot be made unless speciation and bioavailability are adequately addressed. The sediments in Bear Creek have high levels of sulfides, and previous research has suggested that chromium, while highly elevated in the system, should not be present in its toxic form within these sediments (Graham et al., 2009). While assessments of metal partitioning have shown minimal levels of metals in sediment porewater (Hansen et al., 2009), some researchers have put forth alternate explanations for toxicity, positing that exposure of organisms to metals may occur primarily through ingestion of particles, rather than through the porewater (Long et al., 2013). Additionally, redox conditions may have changed throughout the course of the test as oxidation of sediments occurred due to aeration of the beakers, thus increasing bioavailability of metals, which would not

occur in the field. These deeper sediments may be less toxic in situ, as metals in the lower portions of the sediment would be more likely to be reduced than oxidized (lower exposure to air, less biological availability). An analysis of hydrogen sulfide concentrations at the surface and at depth, perhaps conducted without introducing oxygen to the system, could shed light on the bioavailability and toxicity of the metals at the site. Additionally, though high sulfide levels would likely bind metals, hydrogen sulfide itself could be a contributor to sediment toxicity (Sims and Moore, 1995).

Organic contaminants, including PAHs, may be significant drivers of toxicity at the site. PAHs, PCBs and TPH were all found to correlate with mortality, with TPH exceeding PEC levels at all sites throughout the system. The measure of TPH, defined as a family of several hundred compounds originating from crude oil, contains a mixture of all extractable hydrocarbons within the sample, including PAHs and PCBs (ATSDR, 1999). TPH showed a very strong relationship with toxicity, which could be due to a) other organic contaminants not measured by this studied contained within the larger category of TPH, or b) driven by PAHs or PCBs that make up part of the TPH category.

Spatial interpolation was used to predict toxicity and contaminant concentrations throughout the system. The interpolated toxicity results predict a hub of toxicity near the shore of Sparrows Point, with its epicenter at the mouth of Tin Mill Canal, a point of discharge for the Sparrows Point site (EA, 2011). In general, interpolations for metals concentrations predicted a similar distribution and pattern of probable PEC exceedances, with the high contaminant concentrations predicted at the Tin Mill Canal outfall. However, these results also demonstrate an estimated hot spot of metals contamination within Peach Orchard Cove, where minimal toxicity was demonstrated, potentially

related to the existence of boatworks and other marine activities in the area. Though metals are highly elevated in this location, their oxidation states may result in low bioavailability and therefore, low toxicity. Though this area is not known to be toxic to benthic organisms, high concentrations of metals in the cove indicate a potential need for future analysis of risks to ecological and human health, especially in the case of any future site disturbances.

Chapter 2: Application of an Antibody-based PAH Biosensor to Bear Creek Sediments

INTRODUCTION

Polycyclic aromatic hydrocarbons (PAHs) are a group of chemicals composed of carbon and hydrogen, and produced during the incomplete combustion of fossil fuels, wood, or other organic substances, and are found in oil and coal. The PAH class contains more than 100 different compounds, and usually occur as complex mixtures rather than as single compounds (ATSDR, 1995). Of these compounds, 17 are considered as priority compounds, because they are suspected to be more harmful than others, and because they are most highly concentrated at hazardous waste sites (ATSDR, 1995). PAHs can enter the environment through the atmosphere as a result of releases from combustion (fires, volcanoes, automobile exhaust, etc.), as well as entering surface waters from stormwater runoff, as discharge from industrial plants and wastewater treatment plants, and can be released into soils, groundwater or sediments at hazardous waste sites (ATSDR, 1995). Within the environment, PAHs are most likely to be accreted to solid particles (ATSDR, 1995). Thus, sediments in marine and estuarine environments often serve as reservoirs for these compounds.

Bear Creek, near Baltimore, Maryland, is an industrialized site with sediments highly contaminated by PAHs. Bear Creek is a tributary of the Patapsco River and is directly adjacent to the former Bethlehem Steel manufacturing plant, located at Sparrows Point and suspected to be a source of inorganic and organic pollutants in sediments. In a comprehensive assessment of sediment contamination within the Chesapeake Bay, sites in Bear Creek near Sparrows Point also showed high concentrations of total PAHs, as

well as considerable amounts of oil and tar which stymied accurate organic contaminant quantification (Baker et al., 1997). At this site, sediments have also shown toxicity to benthic organisms since testing began over the past several decades, though sediment toxicity has not been causally tied to any particular class of compounds (McGee et al, 1999, 2012; Klosterhaus & Baker, 2006). Because of the abundance of covarying contaminants, correlating toxicity test results with concentrations of any one contaminant category in the sediments has proven inconclusive, thus far.

PAHs, since they are hydrophobic, are found predominantly bound to solid particles within sediments. However, freely dissolved PAHs are the most bioavailable form of the contaminant to aquatic organisms; therefore, concentrations of PAHs within sediment porewater are an important measure of contaminant bioavailability and potential toxicity risk (ter Laak et al., 2006). Assessing the relationship between porewater concentrations of PAHs and sediment toxicity may be key to understanding the cause of toxicity at the Bear Creek site.

Traditionally, detection and measurement of PAHs within sediment porewater are conducted following an involved process of solvent extraction, fractionation, and analysis by high-performance liquid or gas chromatography (Spier et al, 2011). These methods are not only costly and labor-intensive, but they require a large volume of sediment sample in order to obtain sufficient porewater for analysis, particularly when low levels of contaminants are involved (Spier et al., 2011). Portable tools for contaminant detection are enabling more rapid site assessment.

Immunoassays are a rapidly growing category of detection tools for environmental contaminants, particularly PAHs. Several PAH-detection techniques have

been developed, including some commercially available ELISA PAH test kits (Zhang et al., 2012). Immunoassays for environmental PAH offer several advantages, as Spier et al. (2011) summarize: they are fast, easy to use, portable, cost-effective, and highly sensitive, which enables low-level detection without sample extraction. Cross-reactivity often occurs in immunoassays, with an antibody responding to compounds structurally related to the analyte (Zhang et al., 2012). In the case of PAHs, antibodies to specific species also tend to respond to structurally similar PAHs and related compounds, so immunoassays are ideal for assessing total PAH concentrations rather than concentrations of particular substituents (Spier et al., 2011). The major limitation of these methods is that they are only responsive to total PAHs and not to individual compounds. While specificity is lost in this scenario vs. GC analysis, which typically analyzes for a defined subset of constituents and can quantify them separately, PAH biosensors provide other advantages. By assessing total concentrations of compounds and not just those included on a priority list, the biosensor may present a better prediction of actual total sediment PAH levels, and of resultant toxicity, than GC-based analysis.

Drs. Kaattari and Unger at the Virginia Institute of Marine Science in Gloucester Point, VA developed a biosensor based on a KinExA Inline Sensor (Sapidyne Instruments), and using a monoclonal anti-PAH antibody cultured in mice with specificity to 3- to 5-ring PAHs. They demonstrated detection limits as low as 0.3 μg/L and rapid assessment of samples on site in the field. The system showed excellent correlations with GC-MS analysis (Spier et al., 2011).

The objective of this study is to apply the antibody-based biosensor to Bear Creek sediments and investigate the correlations between biosensor results and those obtained

through GC-MS analysis. The study tests the hypothesis that the PAH biosensor will correlate significantly with GC-MS results. This study assesses the correlation of concentrations obtained from the antibody-based measurement technique to total sediment PAH concentrations. By measuring a subset of samples for porewater PAHs (by GC-MS and PAH biosensor) and through whole sediment PAH analysis methods, the study tests the hypothesis that the biosensor will a) perform comparably to GC-MS and b) yield results that correlate as well with total sediment PAHs as those from GC-MS. The study also investigates correlations between total sediment and porewater concentrations and sediment toxicity. By measuring porewater PAH concentrations on both the PAH biosensor and GC-MS, measuring total PAHs in sediments, and examining correlations between these methods and observed mortality, this study tests the hypothesis that the biosensor-derived concentrations will correlate better with mortality to benthic organisms than GC-derived results or total sediment results.

MATERIALS AND METHODS

Site description and sample collection

Sampling sites for toxicity tests were identified in a grid spanning the width of the channel and in a rough north to south gradient. Sites were identified in areas of previously established and/or suspected toxicity. Sediment samples from twenty stations in the region were collected on June 3-4, 2015, according to standard American ASTM and USEPA protocols (ASTM, 1994; USEPA, 1995). On June 11, two more sites were

sampled. Sediment samples were collected using a full-size Ponar grab sampler and boat-mounted davit. The top two centimeters of multiple Ponar grabs were homogenized to generate testing material, by mixing until homogeneous in texture and color. Sufficient material was collected for both toxicity tests and chemical analyses of sediments. Subsamples for chemical analyses were stored in 250-mL certified amber jars and pre-cleaned Mason jars.

Sediments were sampled at depth by collecting 80 cm cores. Core samples were collected from a subset of six sites (C1, C2, C3, C4, F1 and G1) on July 28th, 2015 and from three further sites (F2, G2, G3) on September 17th, 2015. Core sites were selected based on results from surface toxicity tests, described in chapter 1. Sediment cores were collected using an in-house fabricated coring device, which was operated using a boat-mounted davit. The coring device was comprised of a 4-inch diameter PVC tube connected to adjustable lengths of 2-inch PVC pipe, used to plunge the core through sediment layers. Inside the PVC system, adjustable lengths of iron rods were attached to a rubber valve to create suction within the column, which held collected samples in place.

Cores were capped for transport to the University of Maryland, (College Park, MD) and then were stored vertically in a 4^{o}C refrigerator until processing. Before the initiation of sediment tests, cores were removed from the refrigerator and split open length-wise using sheet-metal sheers. Cores were segmented by depth, with ten-centimeter increments from 0 cm to 40 cm and twenty-centimeter increments between 40 and 80 cm. Subsections of cores were stored in Ziploc® bags in a refrigerator at 4^{o}C. for toxicity testing, and in HDPE tubes for chemical analysis. Repeat cores from sites C1, C3

and G1 were collected on January 9th, 2016 to obtain sufficient volume for both biosensor and GC-MS analysis of sediment porewater.

Biosensor analysis of PAHs

Subsets of surface samples and cores samples were spun at 3,500 rpm for 50 minutes in a centrifuge to separate porewater, which was decanted and frozen until immediately prior to analysis. The day prior to biosensor analysis, polymethylmethacrylate beads (Sapidyne) were coated with the antigen, a pyrene-butyric acid-bovine serum albumin conjugate (PBA-BSA). 200 µg of a 1.68 mg/mL stock was added to 200 mg of beads. The antibody used for experiment was 2G8, an anti-pyrene-butyric acid monoclonal antibody produced within mice, sensitive to three- to five-ring PAHs (Li et al., 2016). The antibody was frozen down and then fluorescently tagged with AlexaFluor 647. The antibody solution was made with 30 mL diluent: 240 µL mAb.

Antibody solution and antigen-coated beads were loaded into the instrument, and voltages were stabilized using water samples. Spier et al (2011) describe the automated sample-handling program of the KinExA Inline Sensor. Upon sample introduction, a coated bead pack (approximately 400 µL) was loaded into the flow cell. Next, 400 µL of fluorescently-tagged antibody was loaded, along with 400 µL of water, standard or sample. After mixing within a mixing syringe, half the solution was discarded while the rest was introduced to the bead back loaded into the flow cell, and fluorescence was recorded by the instrument. As Spier et al. (2011) write, the fluorescence signal "was based on competitive exclusion by free PAHs in the sample such that the amount of [antibody] bound to the antigen-coated beads within the flow cell was inversely

proportional to the concentration of three- to five-ring PAHs in the sample." Fifty percent dimethylsulfoxide in DI water was used for automated rinsing between samples. Samples were loaded 7 at a time.

Standards were prepared with phenanthrene in methanol spiked into DI water at concentrations of 0.5 µg/L, 1 µg/L, 1.5 µg/L, 2.0 µg/L and 2.5 µg/L. Standards were used to create a calibration curve to determine porewater concentrations. After stabilization with water and analysis of standards, samples were introduced. Prior to introduction, samples were filtered through a 0.45 µm filter (Millipore). To bring samples within the calibration range of the instrument, many samples were diluted from between a 1:1 dilution and a 100-fold dilution with deionized water.

GC-MS analysis of PAHs in porewater

Subsets of surface samples and cores samples were spun at 3,500 rpm for 50 minutes in a centrifuge, to separate porewater. Due to high levels of suspended solids within porewater samples, they were filtered prior to GC extraction. Filtration occurred in three phases, with samples first passed through an 8 µm Whatman nucleophore filter, then a 1 µm glass filter, followed by a 0.4 µm Whatman Nucleophore filter. This protocol was modified to use a 0.45 µm HAWP Millipore filter to speed the filtration process.

Between 72 and 200 mL of porewater from each sample was transferred into precleaned separatory funnels, with volumes recorded beforehand. Each sample was spiked with surrogate standards containing 2 µg/mL deuterated PAH surrogate (d-4 dichlorobenzene, d8-napthalene, d10-acenapthene, d-10 phenanthrene, d-12 chrysene, d12-perylene,and 2.05 µg/mL 1,1'-binaphthyl. Each sample was extracted three times

using 20 mL of dichloromethane. For each extraction, 20 mL were added to each separatory funnel, and they were shaken vigorously for two minutes before phases were left to separate and the organic phase was drained from the funnels. The set of extractions contained both a laboratory blank and a matrix spike. The volume of the samples was reduced under a gentle stream of nitrogen using a TurboVap® evaporator (Zymark Corp., Hopkinton, MA, USA) and the internal standard p-terphenyl (ChemService, West Chester, PA, USA) was added. The extracts were analyzed either on a Varian 3400 Gas Chromatograph (GC) using a Varian CP-8200 Autosampler coupled to a Saturn 4D GC/MS/MS or a Varian CP-3800 GC using a CP-8400 autosampler coupled to a Saturn 2000 GC/MS/MS ion trap MS (Varian Inc., Walnut Creek, CA, USA) both operated in electron ionization (EI) mode (70 eV). Both were equipped with a split/splitless injectors maintained at 320°C. The carrier gas was He and injections were made in splitless mode on a DB5, 60 m x 0.32 mm x 0.25 μm film thickness capillary columns from J&W Scientific (Folsom, CA, USA). The GC temperature program was 75°C to 350°C at 4 °C/min with an initial hold of 1 min, final hold time was 1.25 min. The Saturn 4D GC/MS manifold and the transfer line temperatures were 270°C and 315°C, respectively. The Saturn 2000 GC/MS trap, manifold and transfer line temperatures were 245 C, 80 C and 320 C respectively. Scans were 100 to 500 m/z for 6 to 71 min; selected ions were used to quantify the targeted analytes. A seven to ten point calibration curve was used for the analyses of individual analytes with either the Varian MS Workstation software package, version 5.2 or 6.8 (Varian). The limit of detection was approximately 0.01 μg/l per analyte on both instruments. A laboratory blank was processed with each set of water samples (Li et al., 2016).

GC-MS analysis of PAHs in total sediments

Analysis of ten whole sediment sub-samples was performed for PAHs at Texas A&M University. Samples H1, G1, G3, F1, E1, E2, D1, D2, C1, and C3 were mailed to Dr. Terry L. Wade at Texas A&M using appropriate handling and chain of custody procedures. The sediments were extracted for total PAH, PCB and TPH analysis using methods previously described by Kirman et al (2016). For TPH and PAH measurement, a flame ionization detector was used for quantitative detection. For PCBs, GC-MS in selected ion mode (SIM) was used (Wade, personal communication, 2016).

Statistical Analysis

To assess relationships between PAH concentrations in surface sediments and cores, and mortality of sediments during toxicity tests, both linear and logarithmic correlations were examined and compared.

RESULTS

GC-MS analysis vs. biosensor analysis of PAHs

A subset of eight surface sites were analyzed on both the PAH biosensor and using GC-MS. Of these eight samples, one experienced leaking during the extraction process in preparation for the GC run. In light of its very low recovery value, this sample was excluded from the analysis. Eight additional samples from core segments were assessed with both methods.

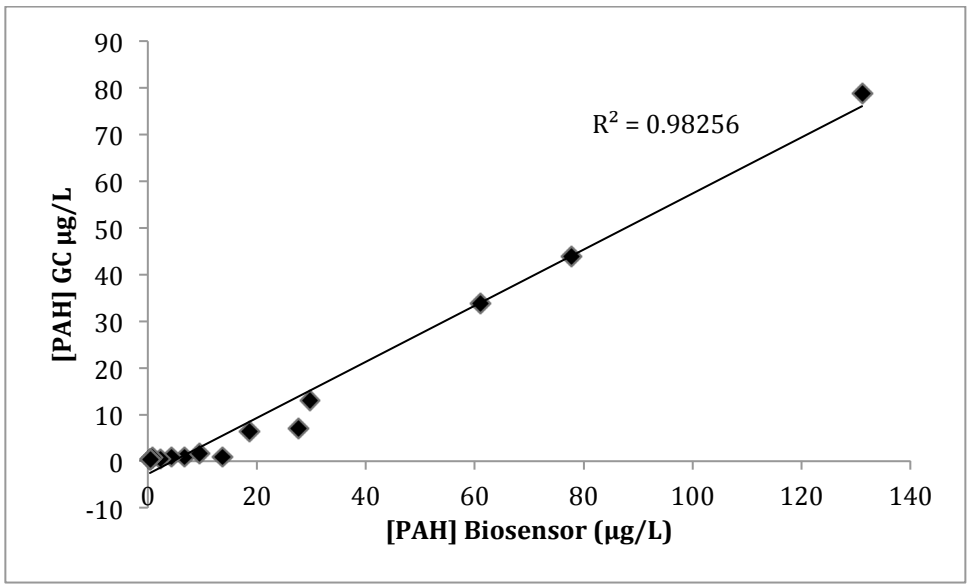


Figure 2.1. Biosensor PAH concentrations and PAH concentrations derived from GC-MS

The correlation between biosensor PAH concentrations and GC PAH concentrations yielded an R^2 value of 0.98, showing excellent correlation between the two methods. In general, the biosensor measured higher concentrations than GC-MS, leading to a slope of less than one.

Table 2.1. PAH values at a subset of surface and core sites, measured by both GC-MS and PAH Biosensor

Site	[PAH] GC (µg/L)	[PAH] Biosensor (µg/L)
H1	0.9	6.7
G2	0.9	4.4
F1	0.5	2.3
E2	0.7	0.8
D1	0.5	0.5
C1	0.4	0.4
B1	0.4	0.5
C1 0-10 cm	0.9	13.7

C1 10-20 cm	6.4	18.6
C1 20-30 cm	13.0	29.8
C1 30-40 cm	33.8	61.0
C1 40-60 cm	43.9	77.7
C1 60-80 cm	78.8	131.1
G1 40-60 cm	1.7	9.5
G1 60-80 cm	7.0	27.6

PAH concentrations in porewater, surface and cores

PAH concentrations ranged from 0.3 µg/L to 6.72 µg/L in surface sediments. PAH values were low in general within surface sediment porewater, but highest concentrations were found in near-shore areas, particularly near the Tin Mill Canal outlet.

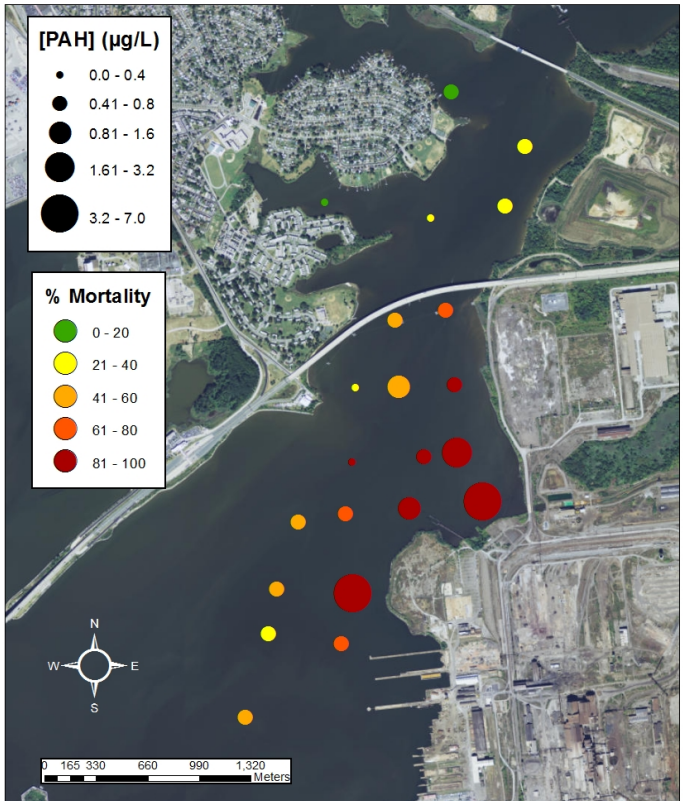
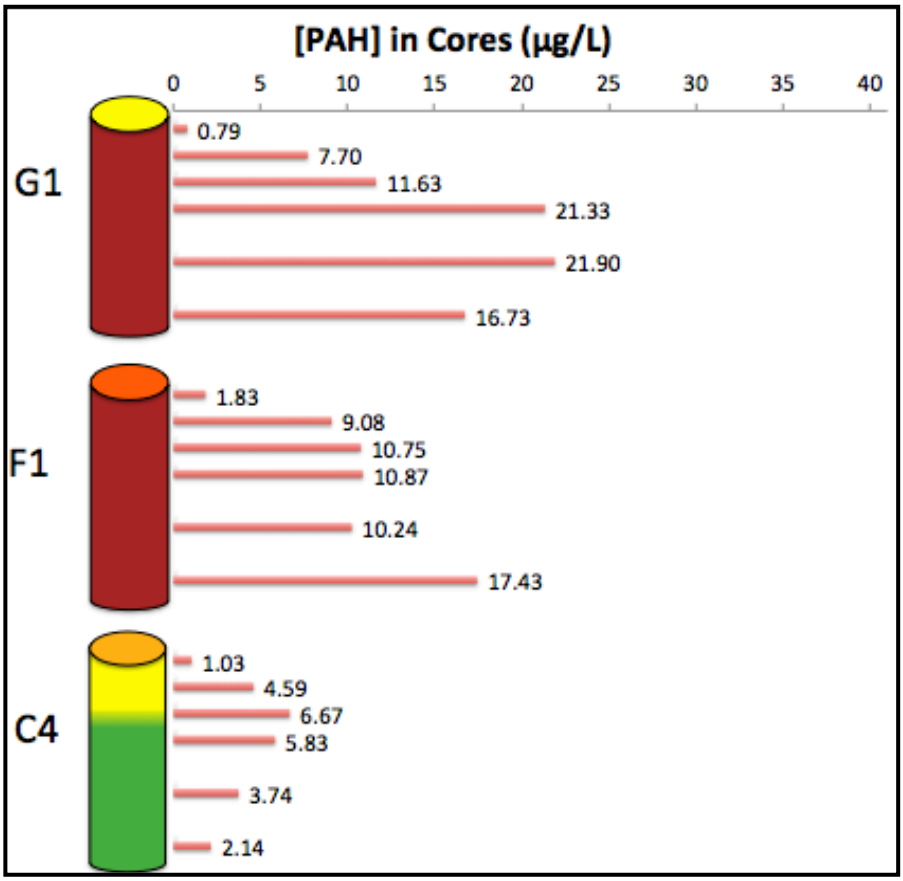
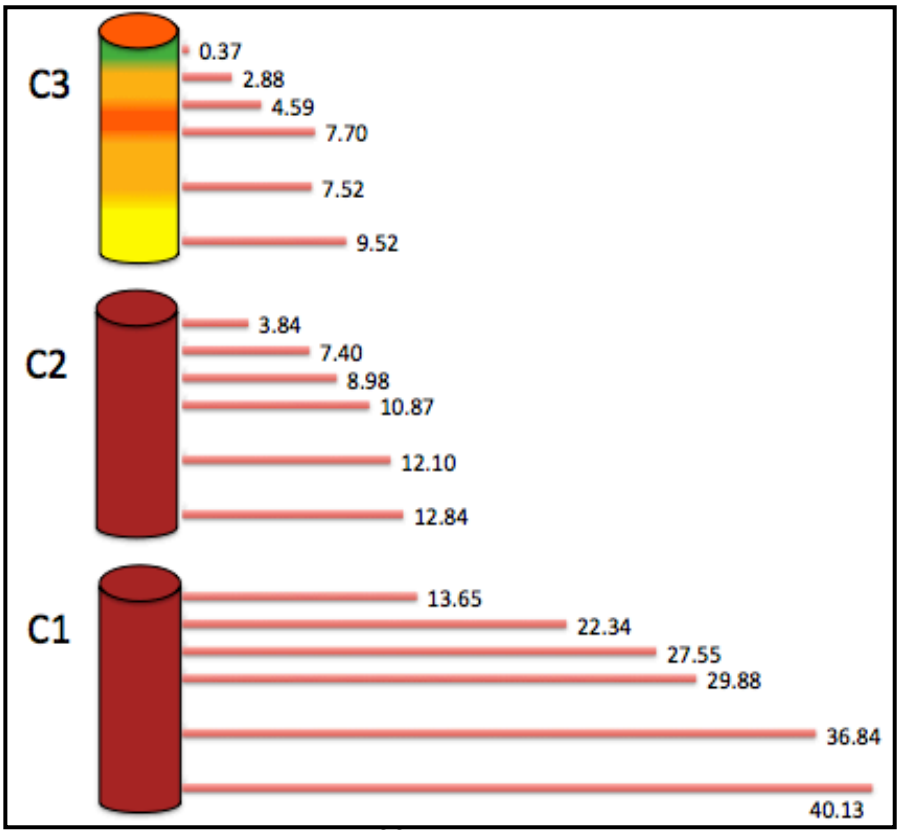


Figure 2.2. Spatial display of PAH and sediment toxicity results


[PAH] in Cores (µg/L)
0
5
10
15
20
25
30
35
40
G1
0.79
7.70
11.63
21.33
21.90
16.73
F1
1.83
9.08
10.75
10.87
10.24
17.43
C4
1.03
4.59
6.67
5.83
3.74
2.14


C3
0.37
2.88
4.59
7.70
7.52
9.52
C2
3.84
7.40
8.98
10.87
12.10
12.84
C1
13.65
22.34
27.55
29.88
36.84
40.13

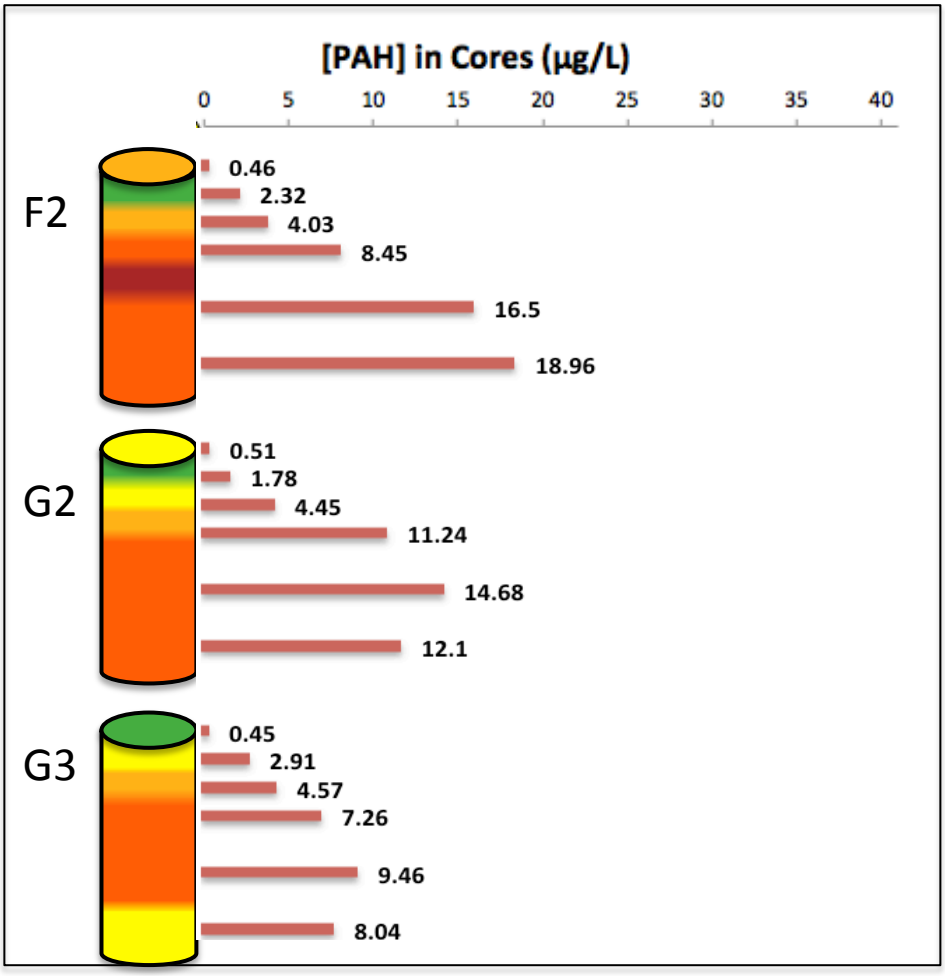


Figure 2.3. PAH Concentrations within core segments

In general, PAH concentrations increased with depth in sediment cores, with 60-80 cm segments exceeding the 0-10 cm segments by up to 27 times. Core C4 shows a unique distribution, with higher concentrations of PAHs in the top core segments vs. the lower sections. As Figure 2.3 shows, high mortality generally corresponded with high concentrations of PAHs within core segments.

In January of 2016, a second set of cores from sites G1, C1 and C3 were collected, and assessed with both toxicity tests and the PAH biosensor measurements of total PAHs in porewater. Cores C1 and C3 showed consistent distributions, while profound differences are seen in Core G1. Substantial differences were observed in Core C1, with a January PAH concentration of 131.10 µg/L in January compared to 40.13 µg/L in August. Concentrations in the lower portions of cores C3 and G1 also increased. Both toxicity and PAH data suggest spatial and/or temporal variability at the study site.

Table 2.2. Comparison of two core toxicity tests run in August, 2015 and January, 2016.

Core Section	**PAH Concentration (µg/L), Original**	**Mean % Survival Original**	**PAH Concentration (µg/L), New**	**Mean % Survival New**
C1 0-10	13.65	0	13.68	5
C1 10-20	22.34	0	18.63	3
C1 20-30	27.55	0	29.80	0
C1 30-40	29.88	0	61.03	0
C1 40-60	36.84	0	77.70	0
C1 60-80	40.13	0	131.10	0
C3 0-10	0.37	85	0.57	71
C3 10-20	2.88	51	3.82	50
C3 20-30	4.59	45	9.26	51
C3 30-40	7.70	25	11.31	33
C3 40-60	7.52	51	11.40	46
C3 60-80	9.52	60	11.88	43
G1 0-10	0.79	18	1.41	70
G1 10-20	7.70	3	5.81	36
G1 20-30	11.63	0	10.96	24
G1 30-40	21.33	0	8.80	11
G1 40-60	21.90	1	9.53	25
G1 60-80	16.73	3	27.55	46

Correlations of mortality and PAH concentrations in porewater and total sediment

Toxicity test data was compared to three different measures of PAH concentrations within Bear Creek sediments: Porewater concentrations obtained through biosensor and GC analysis in this study, and PAH concentrations in total sediment, analyzed for a subset of sites at Texas A&M (Wade Laboratory). Porewater and total sediment PAHs were measured for 6 sites on split samples, which were also subjected to toxicity testing. Correlations between porewater concentrations of PAHs and total sediment PAHs were assessed for both porewater methods (GC and Biosensor).

Table 2.3: PAH concentrations in sediment and porewater measured by three methods

Site	**$[PAH]_{pw}$ (Biosensor)**	**$[PAH]_{pw}$ (GC)**	**$[PAH]_{total}$**	**Mortality**
H1	0.47	0.4	12.3	24
G2	0.39	0.4	16.3	33
F1	0.51	0.5	15.3	72
E2	0.83	0.7	18.4	58
D1	2.32	0.5	45	100
C1	4.37	0.9	49.9	99

Six sample sites (H1, G2, F1, E2, D1 and C1) were analyzed by all three PAH measurement methods. PAH pore water concentrations were in a similar range for both analysis methods, though the biosensor measured slightly higher values than GC in all cases. Sediment PAH concentrations were significantly higher than porewater concentrations, reflecting the partial partitioning of PAHs into porewater.

Correlations were examined between total sediment concentrations of PAHs, and PAHs in porewater measured by each of the two methods. For these relationships, a logarithmic curve was the best fit line for the data distribution.

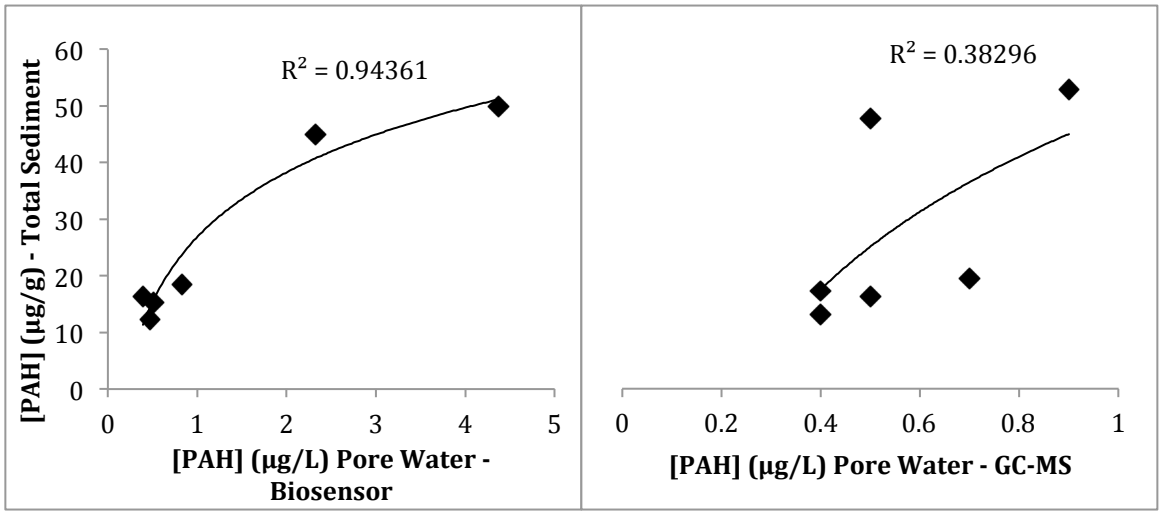


Figure 2.5. Correlations of total sediment PAH concentrations with porewater PAH concentrations as assessed by biosensor (left) and GC-MS (right).

PAH concentrations in porewater as assessed by the biosensor correlated more strongly with total sediment PAH concentrations than those determined with GC-MS, with R^2 values of 0.94 and 0.38, respectively.

Correlations between mortality from sediment toxicity tests and a) total sediment PAHs, b) porewater PAHs measured by GC, and c) porewater PAHs measured by biosensor were assessed. Of the three, total sediment PAHs were correlated most strongly with mortality, with an R^2 value of 0.79. Because the sediments are so highly contaminated with other constituents, the correlation would not be expected to be particularly strong – a caveat that must be considered when interpreting these mortality results.

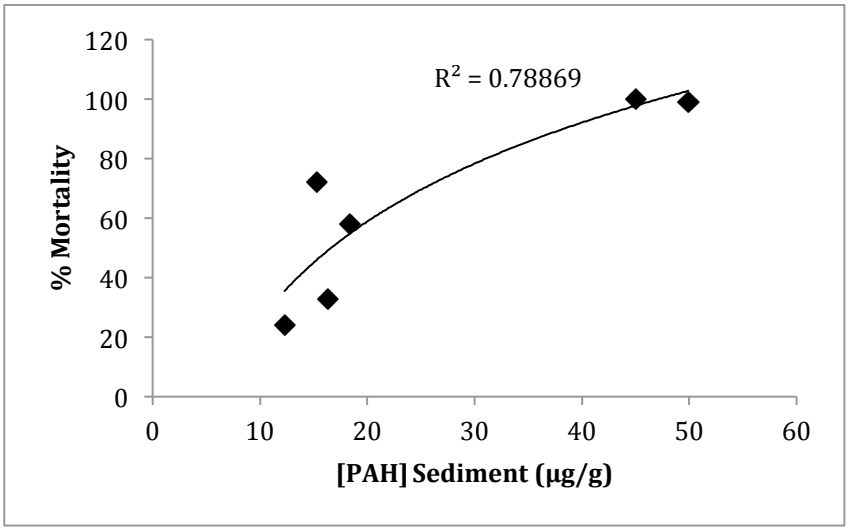


Figure 2.6. Total Sediment PAH concentrations (surface) correlated with % mortality

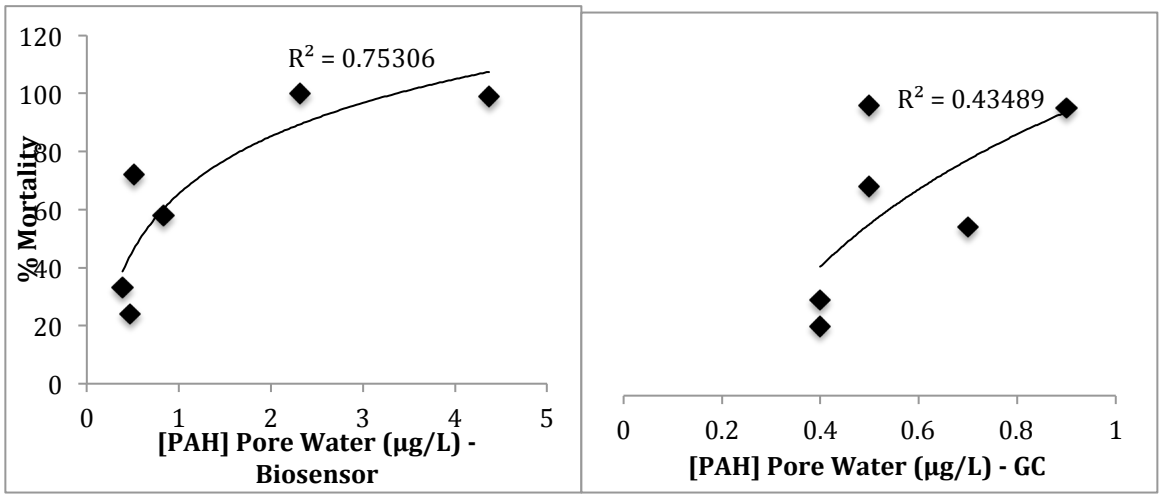


Figure 2.7. PAH concentrations in sediment porewater correlated with % mortality from sediment toxicity tests.

Of the two porewater PAH analysis methods, biosensor concentrations were more strongly correlated with mortality for the six surface sites than GC-MS-derived concentrations.

Correlations were also evaluated for a larger set of samples that included both surface samples and core segments. Complete mortality was observed in sediments

ranging from low to high PAH concentrations, resulting in skewed relationships. Nonetheless, the mortality – PAH concentration relationship was stronger for the biosensor-analyzed samples than for the GC-analyzed set. A logarithmic relationship fit best in all cases when comparing mortality to PAH concentration, indicating a threshold concentration above which mortality reaches 100% and cannot increase any further.

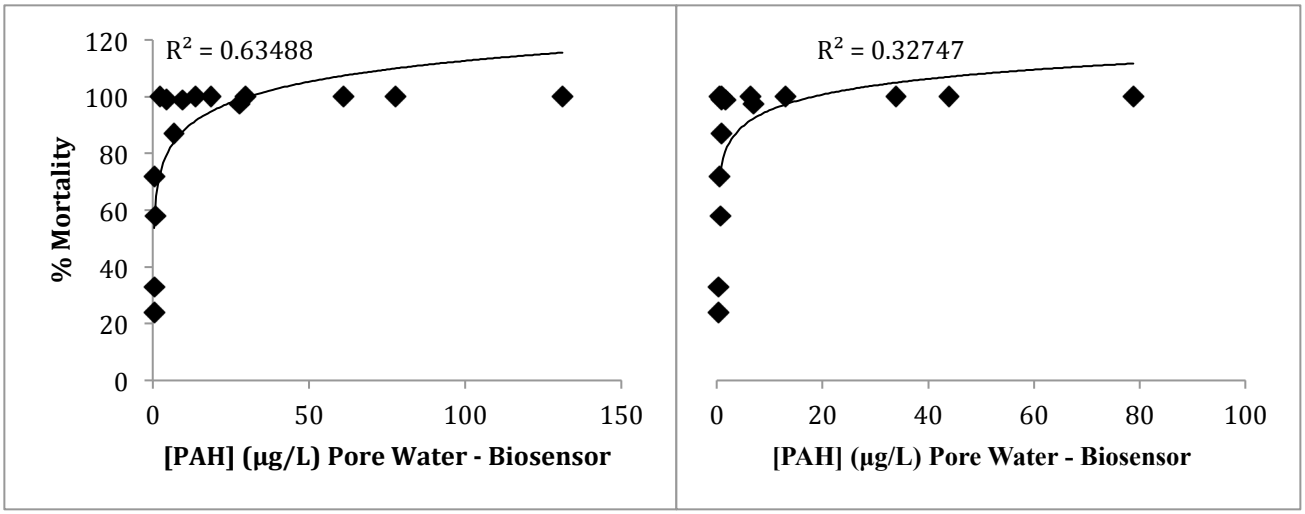


Figure 2.8. Correlations of sediment mortality with PAH concentrations in porewater measured by biosensor (left) and GC (right).

DISCUSSION

.

The comparison of PAH biosensor measurements of total PAHs in porewater to those obtained with GC-MS showed an excellent correlation between the two methods, confirming the study's prediction that the two methods would perform comparably. It is worth noting that the biosensor-derived PAH values were uniformly higher than those obtained from GC-MS. It is possible either that the biosensor overestimates PAHs, or that

it is actually a more accurate representation of total PAH content due to the antibody's sensitivity to the entire suite of 3- to 5-ring PAHs, vs. GC analysis of only a subset. As Spier et al. (2011) write, many studies of immunoassays have reported overestimating PAH concentrations, with this phenomenon occurring through the crossreactivity of the antibody with similar compounds in the sample. However, they hypothesized that traditional analysis may represent an underestimate of the actual total PAH concentration if alkylated species are not considered (Spier et al., 2011).

When compared with total sediment PAH concentrations, the biosensor showed a better correlation than GC-MS measured values. This indicates that the biosensor's higher values may be a better reflection of total sediment concentrations than the lower values obtained by GC-MS, which again supports the hypothesis of comparability between the two methods. Concentrations of PAHs obtained by each of the two porewater methods were assessed in relationship to mortality. Between the two methods, the biosensor showed a stronger correlation with mortality that was statistically significant. PCBs and TPH both showed a significant relationship with mortality. Each of these constituents were compared to toxicologically-based screening levels at which biological effects would be expected in Chapter 1. Whole sediment PAH concentrations exceeded the predicted effects concentration (PEC) and the effects-range median concentration at 3 of the 22 sites, while PCBs were elevated above the PEC at 7 out of 10 sites assessed (Table 1.8). Total petroleum hydrocarbons were uniformly elevated above the SL-SQS, a screening level used to predict adverse effects to benthic organisms. Although PAHs were found to be significantly correlated with mortality, results from

surface assessment of sediments indicates that multiple organic stressors are likely to be involved in causing toxicity within Bear Creek.

Several measured contaminant classes correlated with mortality at the site, but sediment toxicity may also derive from a complex mixture of measured and unmeasured organic contaminants. The measure of TPH, defined as a family of several hundred compounds originating from crude oil, contains a mixture of all extractable hydrocarbons within the sample, including PAHs and PCBs (ATSDR, 1999). TPH showed a very strong relationship with toxicity, which could be due to a) other organic contaminants not measured by this studied contained within the larger category of TPH, or b) driven by PAHs or PCBs that make up part of the TPH category.

Previous studies have commented on the substantial oil and tar content of Bear Creek sediments, which has interfered with accurate quantification of organic contaminants in the past (McGee et al., 1999; Klosterhaus & Baker, 2006). Organic contaminants in non-aqueous phases liquids (NAPLs), a large category of non-aqueous liquids to which oil and tar would contribute, may be a confounding factor in attempting to correlate organic contaminant classes to toxicity at the site. Typically, concentrationsof PAHs in porewater are driven by sediment-water equilibrium partitioning. However, as Hawthorne et al. (2007) discuss, "any significant amount of a NAPL [non-aqueous phase liquid] hydrocarbon phase could change the sediment-water partitioning to liquid-liquid partitioning," leading to a loss of prediction of both porewater concentrations and toxicity. Though PAHs are a contaminant of concern with this site, they are most likely present in a mixture of other inorganic and organic contaminants, which have together resulted in cumulative toxicity.

Though a cause-and-effect relationship between PAHs in sediment porewater and observed toxicity in Bear Creek sediments could not be demonstrated by this study, the biosensor's advantages were clearly demonstrated. The biosensor requires minimal sample preparation beyond centrifuging and filtering, enabling the quick analysis of a larger number of samples than would have been possible by GC-MS. Preparation alone for GC analysis required many days, while 20 surface samples could be prepared and analyzed in one day, and 54 core segments in 2-3 days, with the biosensor. Additionally, the biosensor was able to provide quantitative measurements of PAHs in porewater using a greatly reduced volume. While GC analysis required approximately 100 mL, biosensor analysis could be completed with less than 10 mL of pure samples, and even less of samples that required dilution. This enabled the testing of porewater from sediment cores that were also used for toxicity assays. The small volume of sediment contained within core sections would have prohibited running toxicity assays and chemistry analysis on the same samples. As the repeated analysis of surface samples and core samples demonstrates, significant spatial variability in both toxicity and PAHs can be observed within a site. Thus, the ability to run both toxicity assays and chemical analysis on samples from the same collection event is highly valuable for assessing correlations between mortality and PAH concentrations.

Several proposals for future research were identified throughout the course of the present study. In order to more fully examine correlations between whole sediment PAHs and porewater concentrations derived from the PAH biosensor and GC-MS analysis techniques in Bear Creek sediments, a laboratory-based study using spiked sediment concentrations would be suggested. By spiking sediments, we can reduce or eliminate the

confounding variables presented by other inorganic and organic contaminants that may be potential drivers of toxicity, which would hopefully strengthen observed correlations between PAH concentrations and lethality of sediments to benthic organisms. Additionally, the effect of NAPLs on PAH bioavailability could be assessed within a laboratory environment.

Chapter 3: Application of Handheld XRF to Bear Creek Sediments.

INTRODUCTION

Within the Bear Creek system, concentrations of metals are known to be elevated above predicted effects concentrations (PECs), with zinc and chromium found at particularly high levels. These elements are associated with a long history of industrial activity at the site, and may be in part responsible for demonstrated sediment toxicity. Mason et al. (2004) assessed metals concentrations within sediment cores, and determined that surface concentrations of metals were in many cases lower than those at depth. These high concentrations of contaminants may present ecological or human health risks in this area. However, these risks are still undetermined, in part because of the limited current data set that exists for offshore areas. A more comprehensive assessment of sediment metal concentrations in Bear Creek is important for assessing the extent and depth of contamination at the site, and for guiding management and remediation decisions in the future.

Traditionally, detection of metals in sediments is conducted by laboratory-based methods, including atomic absorption spectrophotometry (AAS) and inductively coupled plasma mass spectrometry (ICP-MS). However, assessing contaminant concentrations in sediments typically requires intensive sample preparation, including digestion of samples with strong acids and oxidizing agents. As Radu and Diamond (2009) describe, the extraction process for heavy metals from soils or sediments involves boiling sediments with concentrated HCl and HNO_3 to completely extract metals, which is still only a pseudo-total extraction. These methods are time- and labor-intensive, and are naturally

destructive to samples, preventing further analysis of samples for other constituents. The ability to perform direct, rapid and non-destructive analysis of metals in soils and sediments is important in efforts to comprehensively map pollution in areas with contaminated sediments. In Bear Creek, for example, metal contamination has only been assessed in a few areas, and to a limited extent within sediment cores (Mason et al, 2004).

Handheld x-ray fluorescence (XRF) is emerging as a technique for assessment of total metals concentrations within environmental samples, including sediment, soil and water (Melquiades and Appoloni, 2004). The Bruker system utilized in this study is an energy-dispersive XRF method. Within the system, x-rays are produced by the bombardment of a rhodium target with accelerated electrons. The produced x-ray beam interacts with atoms in analyzed samples, displacing electrons from the atom's inner orbital shells. As electrons from outer shells fill the vacancies energy is emitted, and the intensity of fluorescence at various wavelengths corresponds to the relative abundance of elements present in the sample. Spectra are produced which may be analyzed to determine the total composition of materials (Drake, 2014). Though the method is not sensitive to different oxidation states and cannot detect elements lower in atomic number than sodium, it is capable of measuring total metals concentrations within samples.

Multiple studies have demonstrated the comparison of handheld XRF with more traditional methods of metals analysis. These studies have shown good correlations between traditional methods, while demonstrates the technique's utility in rapid field assessment. Perroy et al. (2014) analyzed lead contamination at a former trap-shooting range, and were able to assess ten times as many samples as they would have been able to with traditional wet chemistry analytical techniques, contributing to their ability to

develop a fine-resolution three-dimensional map of contamination at the study site. Higueras et al. (2011) comment that handheld XRF may be key to enabling environmental studies in low-technology areas, including developing countries. According to Radu and Diamond (2009), several official methods used by the USEPA and the National Institute for Occupational Safety and Health employ XRF technology. Though laboratory-based methods like ICP-MS and AAS can increase sensitivity and accuracy of assessment, handheld XRF enables faster analysis, allowing real-time pollution mapping of contaminated sites, through which areas for more thorough analysis can be identified (Melquiades and Appoloni, 2004). Additionally, the method is non-destructive, so assessment with XRF can be followed by other methods of analysis on the same sample, while requiring a low volume of sample.

Figure 3.1. Bruker handheld XRF mounted with automatic sample changer

The first objective of this study is to investigate concentrations of various metals in Bear Creek sediment using handheld XRF as a means to increase sample throughput and enable measurement of concentrations in core segments. The second is to validate handheld XRF for analysis of Bear Creek sediments by comparing results from XRF analysis to those obtained through ICP-MS analysis. By comparing results from the two methods, this study tests the hypothesis that handheld XRF will yield comparable results to ICP-MS analysis, and prove to be a useful technique for evaluating a greater number of Bear Creek sediment samples than would be possible with ICP-MS alone.

MATERIALS AND METHODS

Site description and sample collection

Sampling sites were identified in a grid spanning the width of the channel and in a rough north to south gradient. Sites were identified in areas of previously established and/or suspected toxicity. Sediment samples from twenty stations in the region were collected on June 3-4, 2015, according to standard American ASTM and USEPA protocols (ASTM, 1994; USEPA, 1995). On June 11, two more sites were sampled. Sediment samples were collected using a full-size Ponar grab sampler and boat-mounted davit. The top two centimeters of multiple Ponar grabs were homogenized to generate testing material, by mixing until homogeneous in texture and color. Sufficient material was collected for both toxicity tests and chemical analyses of sediments. Subsamples for chemistry analysis by ICP-MS were stored in 250-mL certified amber jars and pre-cleaned Mason jars. Subsamples for metals analysis by handheld XRF were stored in 50-mL conical tubes.

Core samples were collected from a subset of six sites (C1, C2, C3, C4, F1 and G1) on July 28th, 2015 and from three further sites (F2, G2, G3) on September 17th, 2015. Core sites were selected based on results from surface toxicity tests. Sediment cores were collected using a manufactured coring device, which was operated using a boat-mounted davit, described in Chapter 1. Cores were segmented by depth, with ten-centimeter increments from 0 cm to 40 cm and twenty-centimeter increments between 40 and 80 cm. Subsections of cores were stored in Ziploc bags in a refrigerator at 4ºC for toxicity tests, and aliquots were removed for XRF testing and stored in 50-mL conical tubes.

Sample preparation for handheld XRF

Between 10 g and 25 g of each surface and core sample were removed from storage and placed on labeled weigh boats (total mass depended upon available material). Samples were dried overnight in a Fisher Scientific Isotemp Incubator (Model 6550) at 60ºC, and were weighed following drying to enable assessment of dry vs. wet weight. Samples were ground to a fine powder using a ceramic mortar and pestle. All equipment was cleaned between samples with soap and water, nitric acid, acetone, and deionized water to prevent cross-contamination. Powdered samples were covered and stored in polypropylene cups until XRF analysis.

Samples were prepared in several different forms and to several different thicknesses to examine the effect of sample preparation on XRF signal intensity. Four types of samples were prepared: 1) a loose powder in a Prolene® sample cup; 2) a packed powder in a sample cup; 3) a pellet, analyzed without the sample cup; a pellet, analyzed with the sample cup; and a wet sample. Pellets were created using a Fisher Scientific

stainless steel pellet press with a 0.5 inch diameter and a 1.0 inch die cavity. For each of these preparation styles, samples were prepared at thicknesses of 10 mm, 15 mm, and 20 mm, in order to investigate depth to complete signal attenuation.

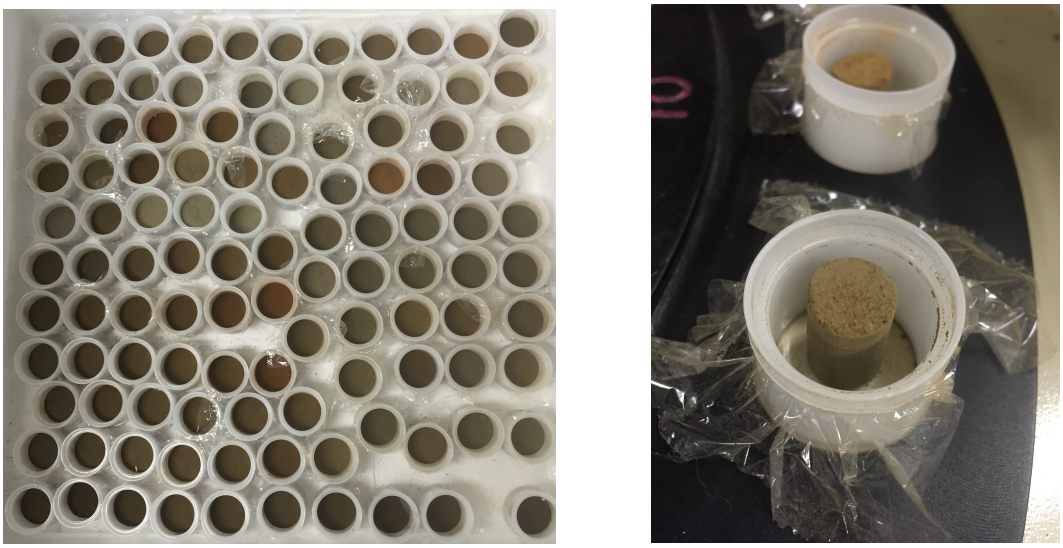


Figure 3.2. Sample cups containing powdered sediments (left) and sediment pellets (right)

Calibration curves were constructed in two different ways. First, a subset of formerly collected samples from within the system were dried and retained to create calibration curves. Calibration curves were created manually using these samples, which were also used to supplement a pre-loaded Mudrock calibration curve provided by Bruker. Another set of calibration standards was created using pure silicon dioxide spiked with multiple metals standard solutions, following a spiking protocol summarized in Table 3.1. All solutions were obtained from Merck-Millipore, and were 1000 µg/mL for the constituent of interest. Standards were created by immersing approximately 7.5 g of pure silicon dioxide in a mixture of metal standards, and then heating beakers at 70^{o} C for a period of two days to evaporate excess liquid and deposit metals on the SiO_2

Table 3.1. Standards were prepared to match the following series of concentrations

	Desired Metal Concentration (µg/g dry weight)					
Standard	**Zinc**	**Chromium**	**Copper**	**Nickel**	**Arsenic**	**Lead**
Blank	0	0	0	0	0	0
1	500	500	100	25	20	100
2	1000	1000	200	50	30	200
3	2000	2000	300	100	40	300
4	3000	3000	400	150	50	400
5	5000	5000	500	200	60	500

Following evaporation, silicon dioxide spikes were mixed with a spatula and placed in sample cups for XRF analysis.

Handheld XRF analysis of sediments

Ground and dried samples were re-homogenized and placed in XRF sample cups (Chemplex industries) constructed with a Prolene® film bottom. Each sample and standard was loaded into a cup to a uniform depth of 10mm. Sample cups were loaded onto the Dewitt sample-changer, which accommodates 20 samples at a time for analysis.

Figure 3.3. Samples loaded in Dewitt sample changer

Samples were analyzed for heavy metals using a Bruker Tracer III SD and affiliated software – S1PXRF, ARTAX, and Dewitt sample-changer software. A non-gridded Prolene® window was applied to the nose of the Bruker Tracer III SD handheld XRF device. The yellow filter (25 um Ti/300 um Al) was inserted into the device using forceps. The Bruker instrument was then mounted beneath the Dewitt sample changer, the instrument was switched on, and communication was established between the instrument and computer software. Within the setup dialogue box, the following options were selected: 2 Bytes per Channel, Accumulation Mode, Advance Header, S1Mode, Number of Channels: 2048. The KTI tube was optimized at 40 keV. Settings used for the analysis of heavy metals were 40 keV and 10-12 uA, with no vacuum. Data was collected in 120 second intervals, with each sample analyzed in triplicate, using the automatic sampling mode on the Dewitt Sample Changer. For spectral analysis, ARTAX software was used. Stored PDZ files were converted to TXT files and spectra were opened in ARTAX. Region of Interest (ROI) analysis was performed on calibration samples for chromium, lead, copper, zinc, nickel, arsenic, and lead, and the settings were saved for future application to surface sample and core sample spectra. For all samples but lead, the K-alpha peak was used; for lead, the L-beta peak was selected because the K-alpha peak overlapped with that of arsenic (peaks are named based on the electron shell from which the electron transition occurs).

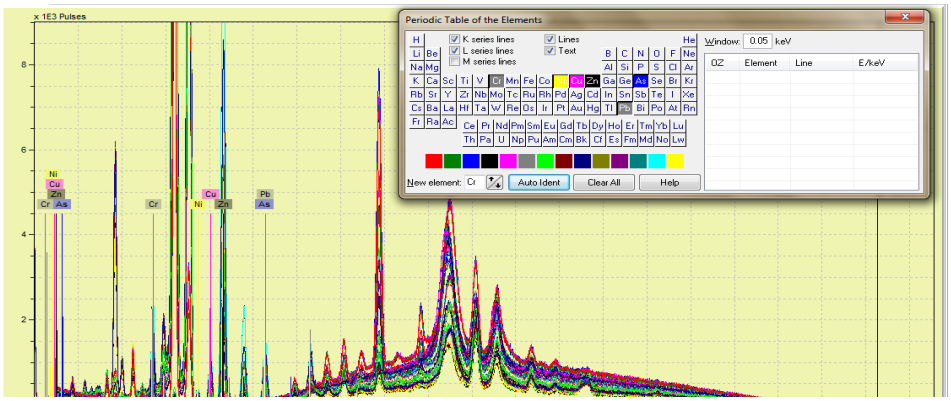


Figure 3.4. Example spectra in ARTAX software, with element selection tool visible.

ICP-MS Analysis of Sediments

Subsamples were retained from each surface sediment sample and core segment for analysis of select heavy metals and organic constituents. Metals analysis was performed at the University of Maryland Chesapeake Biological Laboratory, Solomons, MD. Samples were hand-delivered to Dr. Andrew Heyes after being kept on ice. Inductively coupled plasma mass spectrometry (ICP-MS) was used to analyze concentrations of Cr, Zn, Cu, Ni, Pb, V and As in each sample. Samples were weighed and placed in a VWR Scientific Forced Air Oven at 60^{o}C overnight, and were then reweighed to determine dry weight/wet weight ratios. Another subsample of each sediment was placed in an acid-cleaned quartz flask for microwave digestion, using EPA Method 3052. The Milestone EOTHO-EZ uses quartz reaction vessels placed inside PTFD flasks, which are sealed during the digestion. For digestion, 1-2 g of sediment were placed in the vessel with 9 mL of concentrated ultra pure HNO_3 and 2 ml of concentrated ultrapure HCL. The vessel was capped with a loose fitting quartz cap, and placed in the Teflon flask; 5 ml of 30% H_2O_2 was added to the Teflon sleeve and the sleeve was

sealed. The sample was heated to 180^{o}C in the microwave for 15 minutes then allowed to reflux for 20 minutes. The samples were then cooled and filtered through Whatman No. 41 filter paper by suction filtration and diluted to required volume with deionized water. These extracts were analyzed using a Hewlett-Packard 4500 Inductively Coupled Plasma-Mass Spectrometer (EPA Method 6020). Standard reference materials and Digestion acid blanks were analyzed with each batch - NIST 1646a and NIST 1944 (NY/NJ), and standard additions were done to test for interferences (Heyes, personal communication, 2016).

Statistical Analysis of Data

XRF calibration curves were created for zinc, chromium, copper, nickel, and arsenic by plotting ICP-obtained concentrations of each metal vs. net peak areas in counts per 120 seconds (SiO_2 spikes) The equation of the linear regression from this plot was used to compute concentrations of surface samples analyzed with XRF. Correlations between XRF and ICP-MS were assessed for these surface samples by plotting concentration predicted from XRF calibration curves vs. actual concentrations obtained through ICP-MS. These values were compared to those obtained via modification of Bruker's Mudrock calibration curve. Core sample concentrations were assessed using the Mudrock calibration curve, and concentrations were computed using this adapted calibration curve. To assess correlations between metals and mortality, Pearson Product Moment Correlations were computed for a) surface samples assessed with ICP-MS, b) surface samples assessed with XRF, and c) core samples and surface samples assessed by XRF.

RESULTS

Results – XRF sample preparation and calibration

The impact of sample preparation on XRF signal response was assessed by comparing samples prepared in several different forms and to several different thicknesses. Four types of samples were prepared: 1) a loose powder in a Prolene® sample cup; 2) a packed powder in a sample cup; 3) a pellet, analyzed without the sample cup; a pellet, analyzed with the sample cup; and a wet sample. For each of these preparation styles, samples were prepared at thicknesses of 10 mm, 15 mm, and 20 mm. Zinc intensities, measured in counts/120s, were measured and plotted.

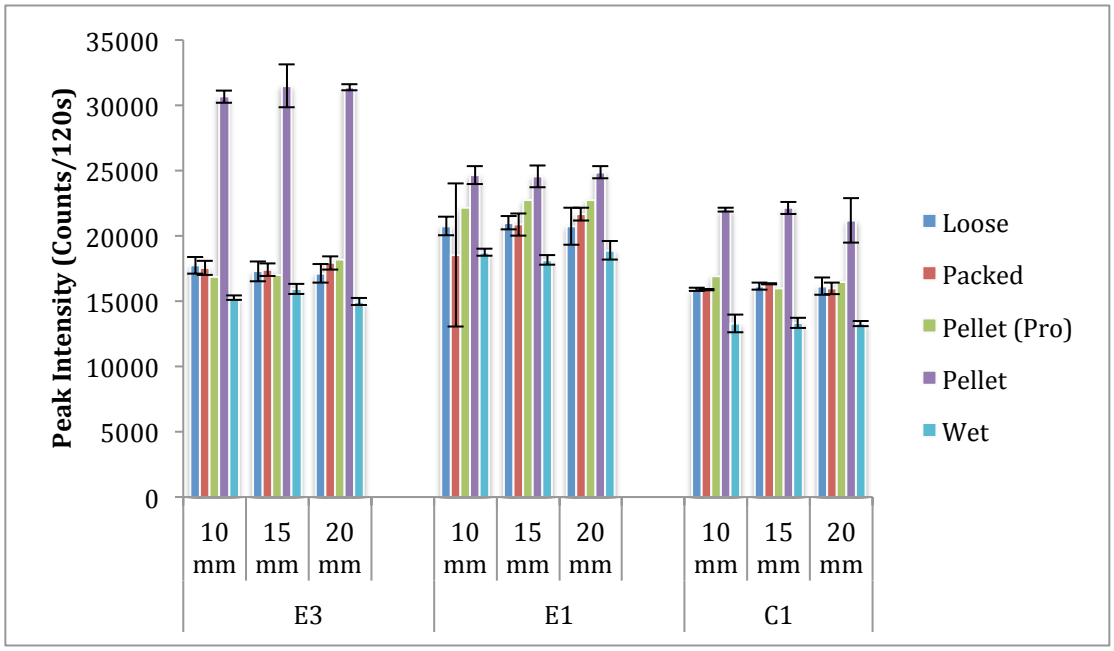


Figure 3.5. Sample preparation comparison for handheld XRF. "Pellet" indicates a pellet run without a sample cup, while Pellet (Pro) indicates a pellet that was placed in a Prolene® sample cup. Raw peak intensities recorded by the XRF software are reported for samples of each style and each thickness. Bars represent an average of three readings, with error bars showing standard deviations of the three readings.

Signal responses were comparable across multiple thicknesses for the pressed powder, loose powder, and pellet forms, when all were assessed within the Prolene® sample cups. Signals were substantially higher when pellets were analyzed without the sample cup. Wet samples were somewhat lower than dry samples, but still within a comparable range.

Several methods were explored for creating calibration curves, including constructing a manual calibration curve out of an old set of samples gathered from within Bear Creek, and using these standards to modify Bruker's built-in Mudrock calibration. Silicon dioxide spiked with metals yielded a superior set of manual calibration curves than in-system samples.

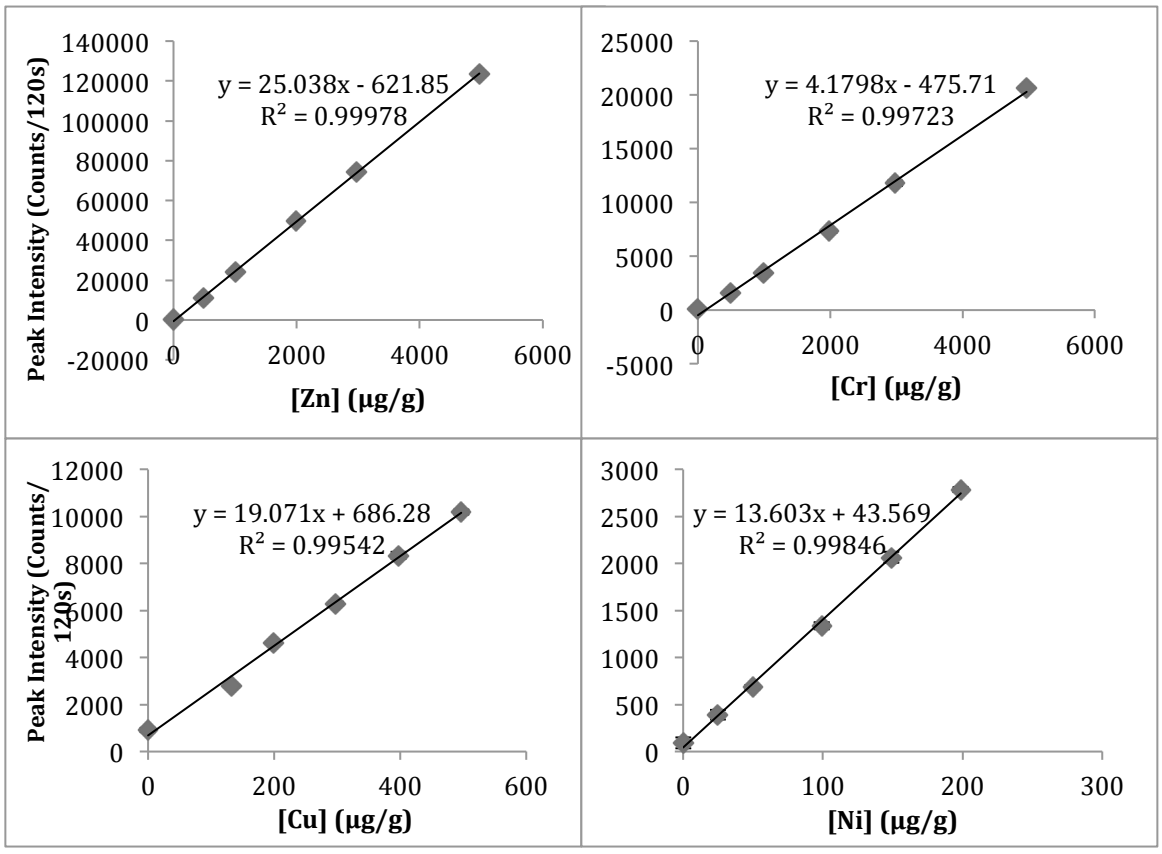


Figure 3.3. Manual calibration curves created with spiked silicon dioxide

Though the SiO_2 spikes yielded excellent metal-peak intensity correlations, when the equations were used to predict metals concentrations in surface samples already validated with ICP-MS, the results showed poor correlations. The standards were not used on their own, but were instead measured and used to supplement Bruker's Mudrock calibration curve, a built-in curve on the XF instrument. The calibration curve can be modified with any number of standards, and applies various inter-element corrections to account for interactions between elements within the spectrum. Correlations were improved for zinc, nickel and copper by using the modified Mudrock curve, though room for improvement remains. Chromium's best correlation was obtained by using the manually-created calibration curve.

Results: ICP-MS vs. handheld XRF

Table 3.2. Concentrations of select metals in Bear Creek sediments, measured by ICP-MS and XRF.

Station	Metal conc. (µg/g)							
	Zn		Cr		Cu		Ni	
	ICP	XRF	ICP	XRF	ICP	XRF	ICP	XRF
H2	1524.70	1647.83	965.60	608.41	196.15	211.03	47.60	102.81
H1	1362.80	1549.40	913.10	662.80	164.20	193.10	58.00	106.24
G1	1513.60	1822.25	987.20	699.88	155.50	192.54	65.60	113.85
G2	1585.70	1643.67	1066.10	712.64	183.30	200.12	60.10	120.79
G3	2229.80	1493.08	1220.10	606.82	315.50	175.19	82.60	96.81
F1	1455.00	1771.66	944.10	770.30	145.00	202.80	51.20	131.21
F2	1389.45	1379.93	931.80	630.66	170.80	169.53	21.65	105.90
E1	1569.20	1809.68	1153.30	939.37	163.80	212.98	63.10	175.48
E2	1168.30	1070.98	853.00	504.26	147.80	157.52	53.40	91.95
E3	920.00	1090.70	547.40	523.72	79.20	102.66	32.60	93.36
D1	2294.90	2316.23	2685.65	1793.40	228.65	269.52	118.25	250.18

D2	1451.40	1561.87	1281.80	947.34	176.20	192.35	71.90	155.55
D3	1145.55	1317.16	882.10	506.73	138.30	186.42	47.85	99.33
C1	1415.95	1649.13	3196.05	1529.99	212.95	293.26	127.70	277.36
C2	1655.30	1824.59	1240.30	753.08	176.60	240.62	76.10	194.72
C3	983.3	1116.41	720.30	544.13	139.60	163.60	59.30	98.96
C4	1378.70	1356.20	799.80	639.51	186.70	198.25	43.50	104.82
B1	2000.90	2356.74	1720.90	1106.60	343.40	470.28	69.20	153.71
B2	986.20	1161.26	856.70	783.46	114.00	161.44	37.10	113.60
A1	705.80	1111.72	601.20	611.28	144.30	235.40	62.40	109.14
A2	604.80	758.82	467.40	405.53	98.50	135.96	38.60	81.94
A'1	774.80	931.41	614.85	537.84	142.45	192.02	43.85	99.53

Handheld XRF results differed substantially from those obtained through ICP-MS analysis, and deviations were different from element to element. For zinc, XRF values were greater than ICP-MS values for 18 out of 22 samples. In sites G3, F2, E2 and C4, XRF values were lower. For chromium, XRF values were lower than those obtained through ICP-MS, except in the case of site A1. For some sites, ICP-MS values for chromium were twice as high as XRF values. Copper XRF values were also generally higher than ICP-MS values, with the same sites deviating from the trend as in the case of zinc. All XRF values for nickel were higher than ICP-MS, sometimes by three times or more. Deviations between ICP-MS and XRF must be evaluated on an element-by-element basis, though copper and zinc do show similar trends to one another.

Correlations between ICP-MS and handheld XRF were explored for the four metals summarized in Table 3.2. For several key metals, measurements from ICP-MS and handheld XRF were reasonably correlated (Figures 3.7, 3.8, 3.9, and 3.10). However, R2 values between .55 and .88 show that calibration curves may require further modifications to correct for differences between samples.

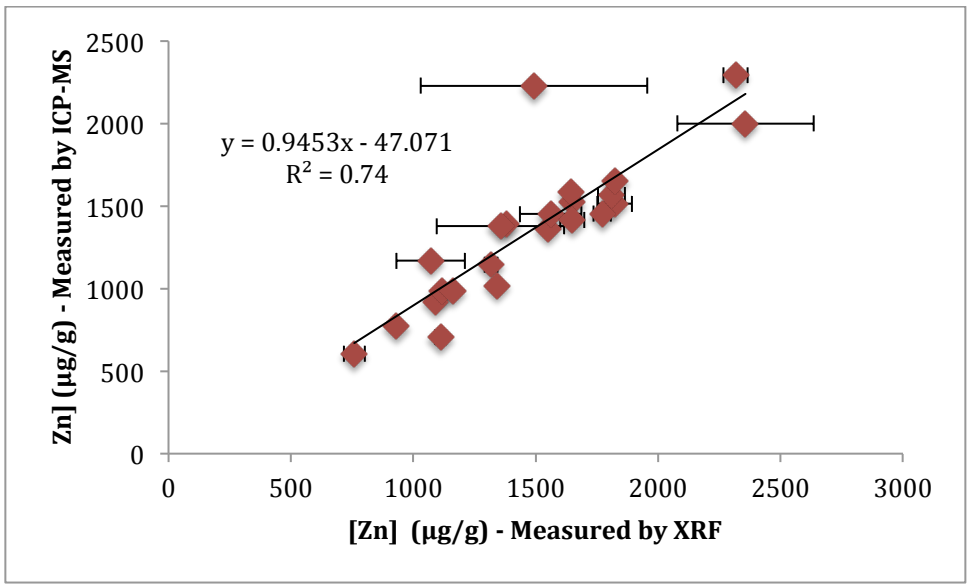


Figure 3.7. Correlation between zinc values as measured by ICP-MS, and as measured by handheld XRF, using the modified Mudrock calibration curve. Each point is an average of three measurements, with error bars representing the standard deviation of the three measurements.

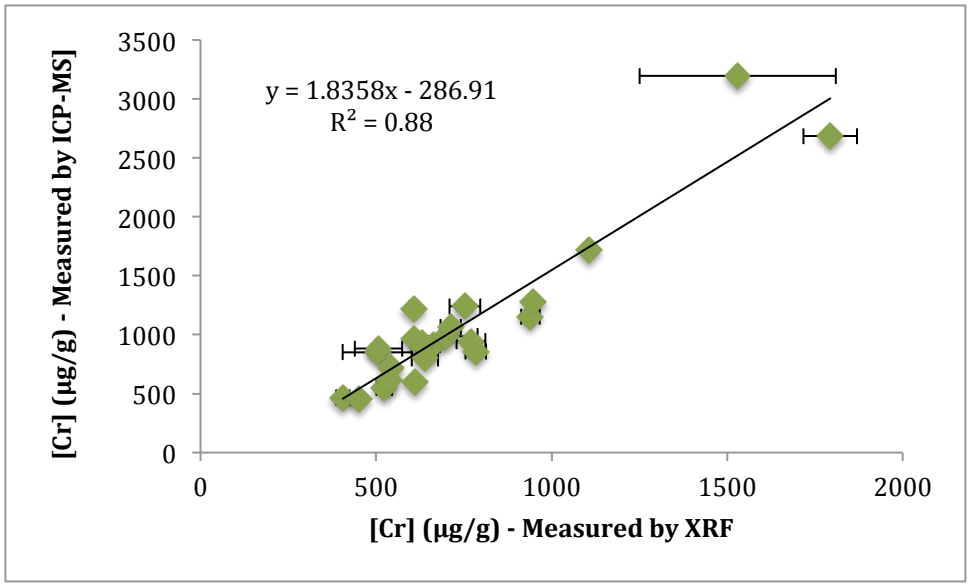


Figure 3.8. Correlation between chromium values as measured by ICP-MS, and as measured by handheld XRF, using a manual calibration curve. Each point is an average of three measurements, with error bars representing the standard deviation of the three measurements.

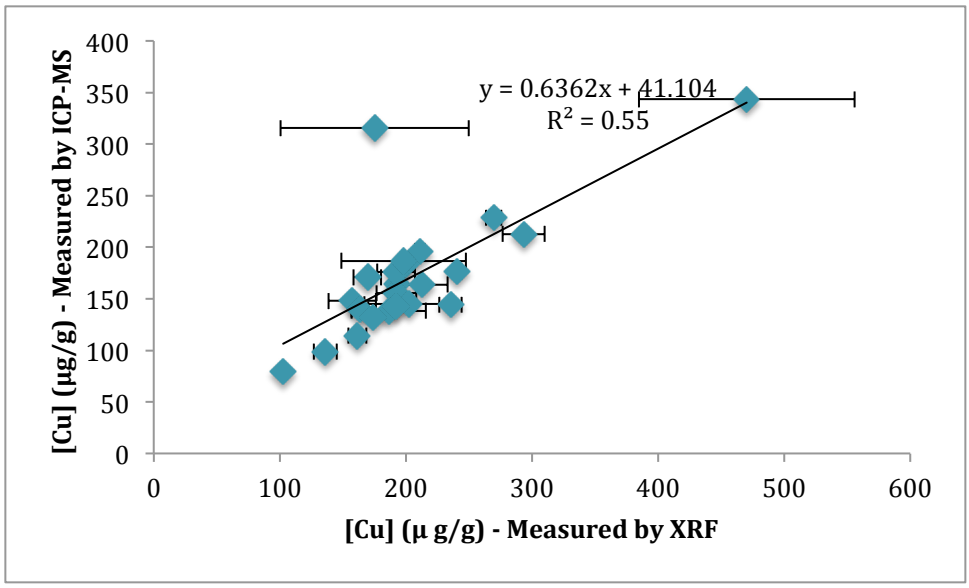


Figure 3.9. Correlation between copper values as measured by ICP-MS, and as measured by handheld XRF, using the modified Mudrock calibration curve. Each point is an average of three measurements, with error bars representing the standard deviation of the three measurements.

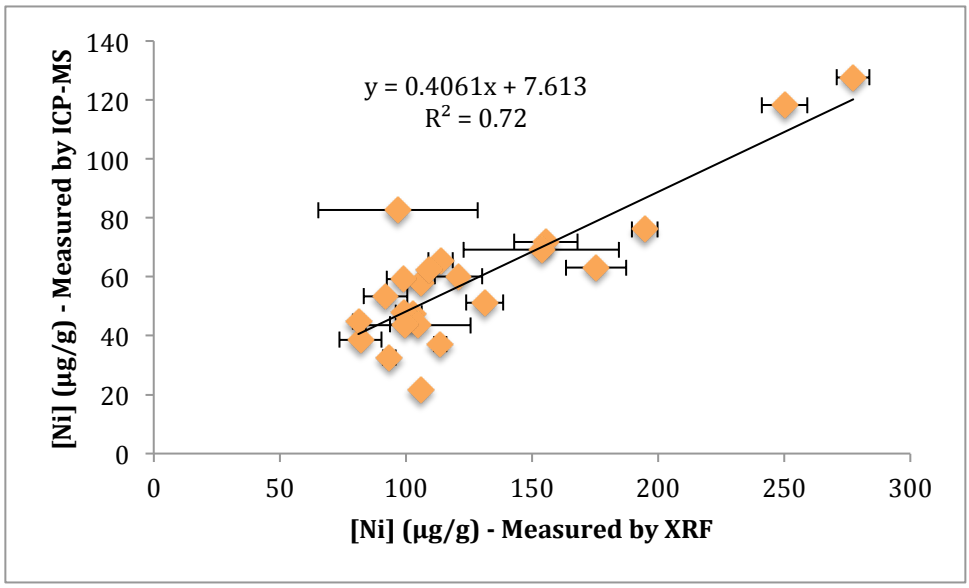


Figure 3.10. Correlation between nickel values as measured by ICP-MS, and as measured by handheld XRF, using the modified Mudrock calibration curve. Each point is an average of three measurements, with error bars representing the standard deviation of the three measurements.

XRF concentrations in sediment cores

Concentrations in core segments were predicted using the modified Mudrock calibration curve for zinc, nickel and copper, and the manual calibration curve for chromium. Though exact values are likely inaccurate due to the need for further calibration curve modifications, relative abundances of elements within cores can be estimated.

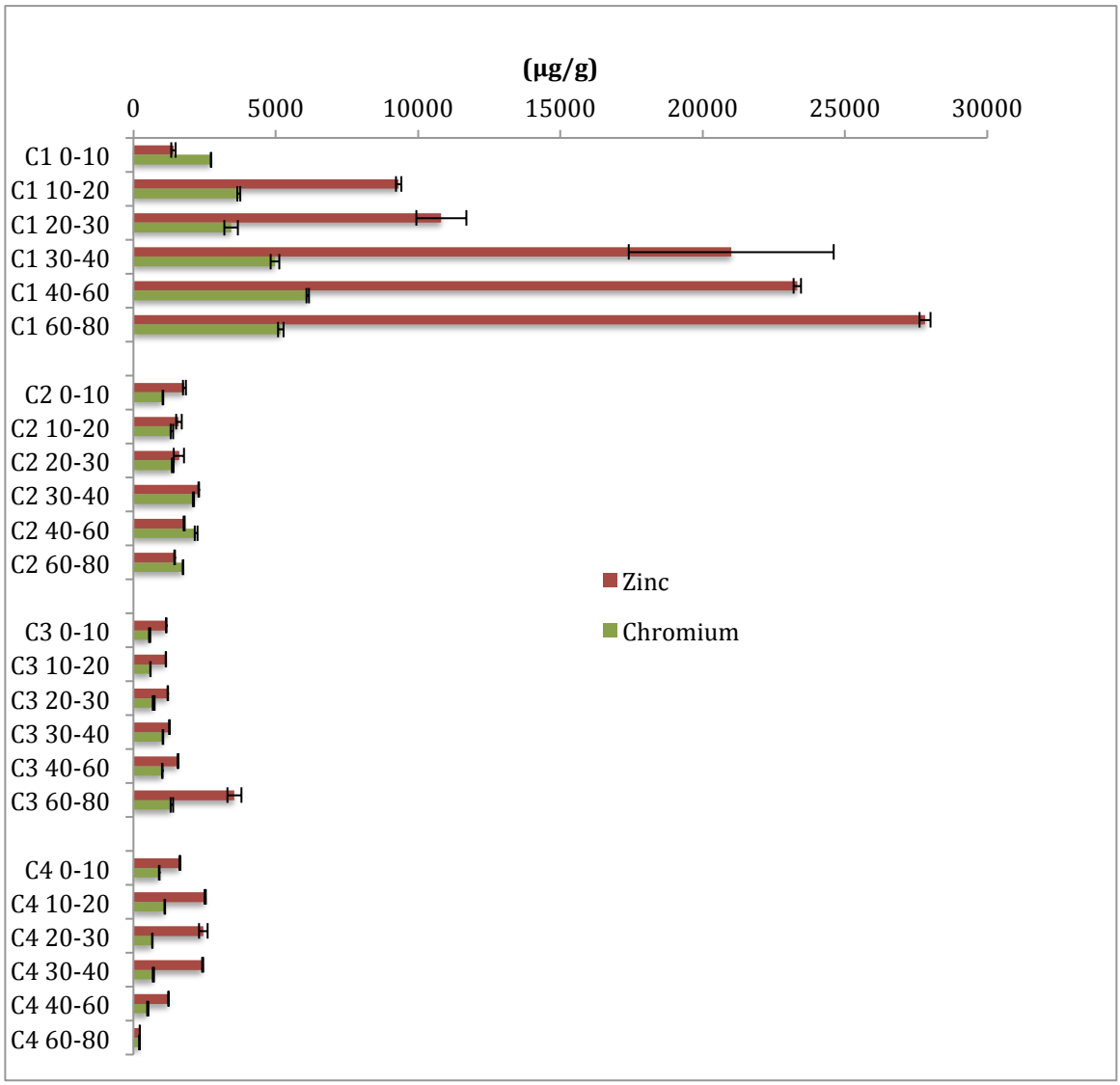


Figure 3.11. Predicted concentrations of chromium and zinc in C-transect cores.

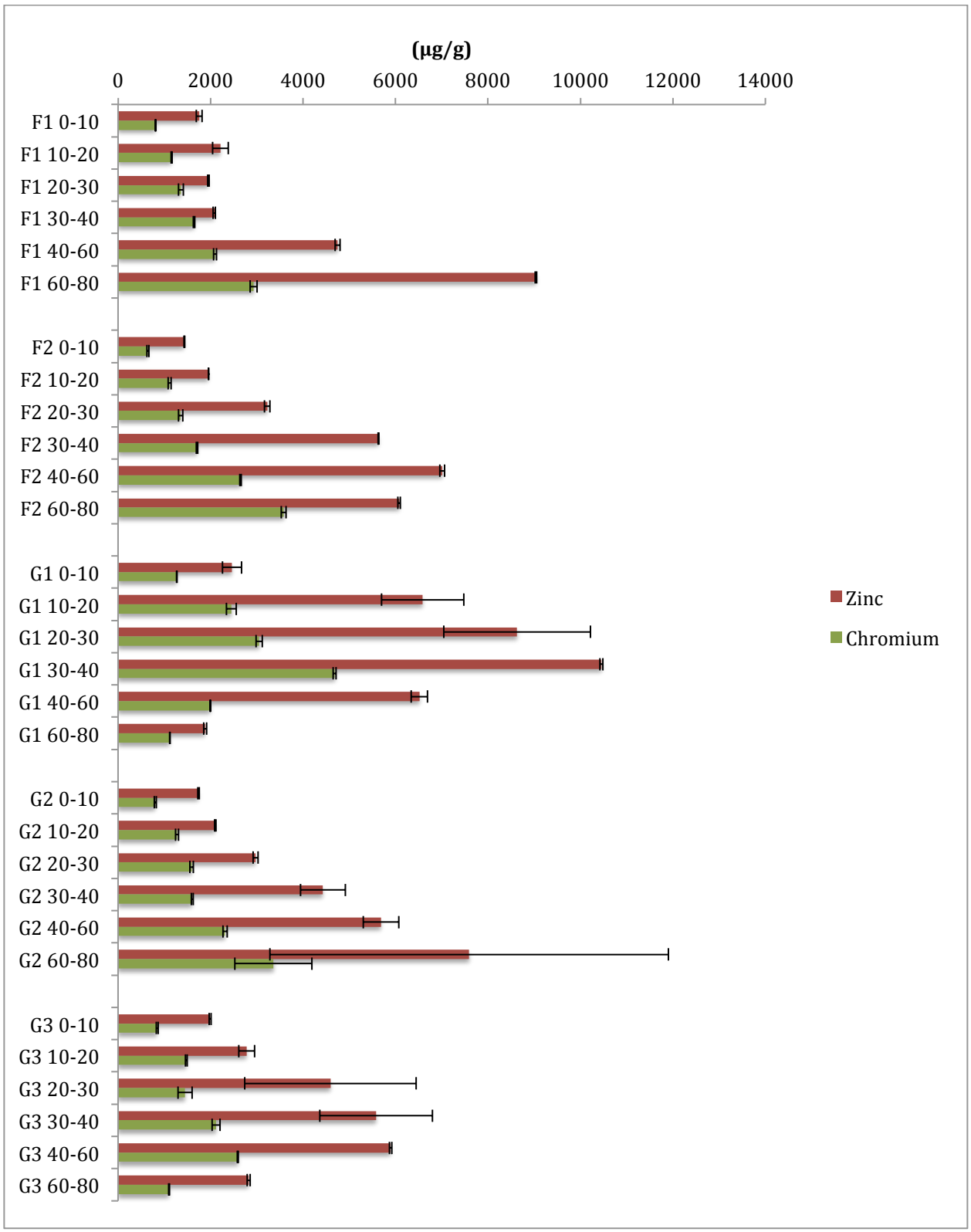


Figure 3.12. Predicted concentrations of chromium and zinc in F- and G-transect cores

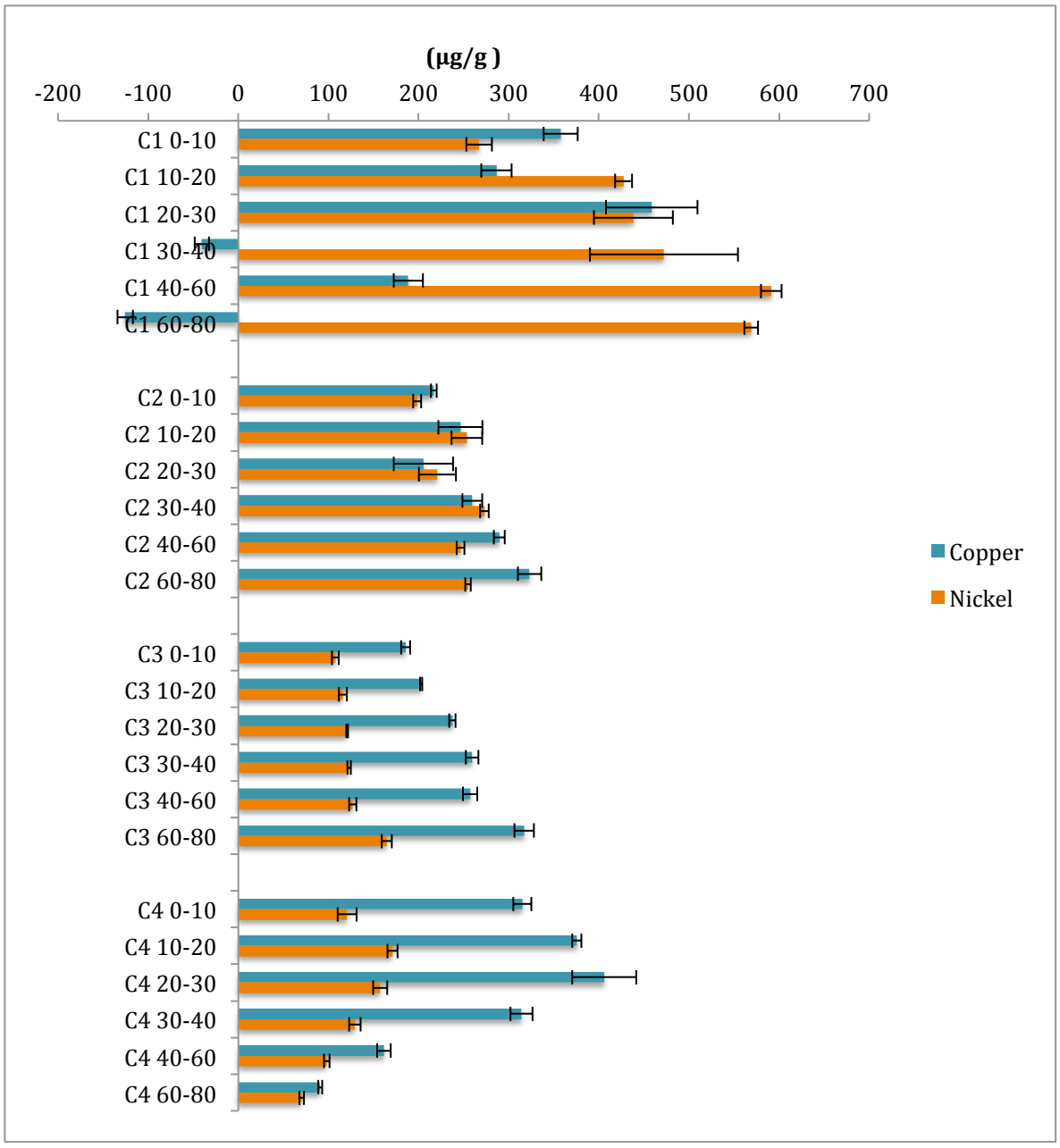


Figure 3.13. Predicted concentrations of copper and nickel in C-transect cores

In several instances, copper was predicted as a negative concentration, indicating errors with the calibration curve that may be impacting XRF measures of this element.

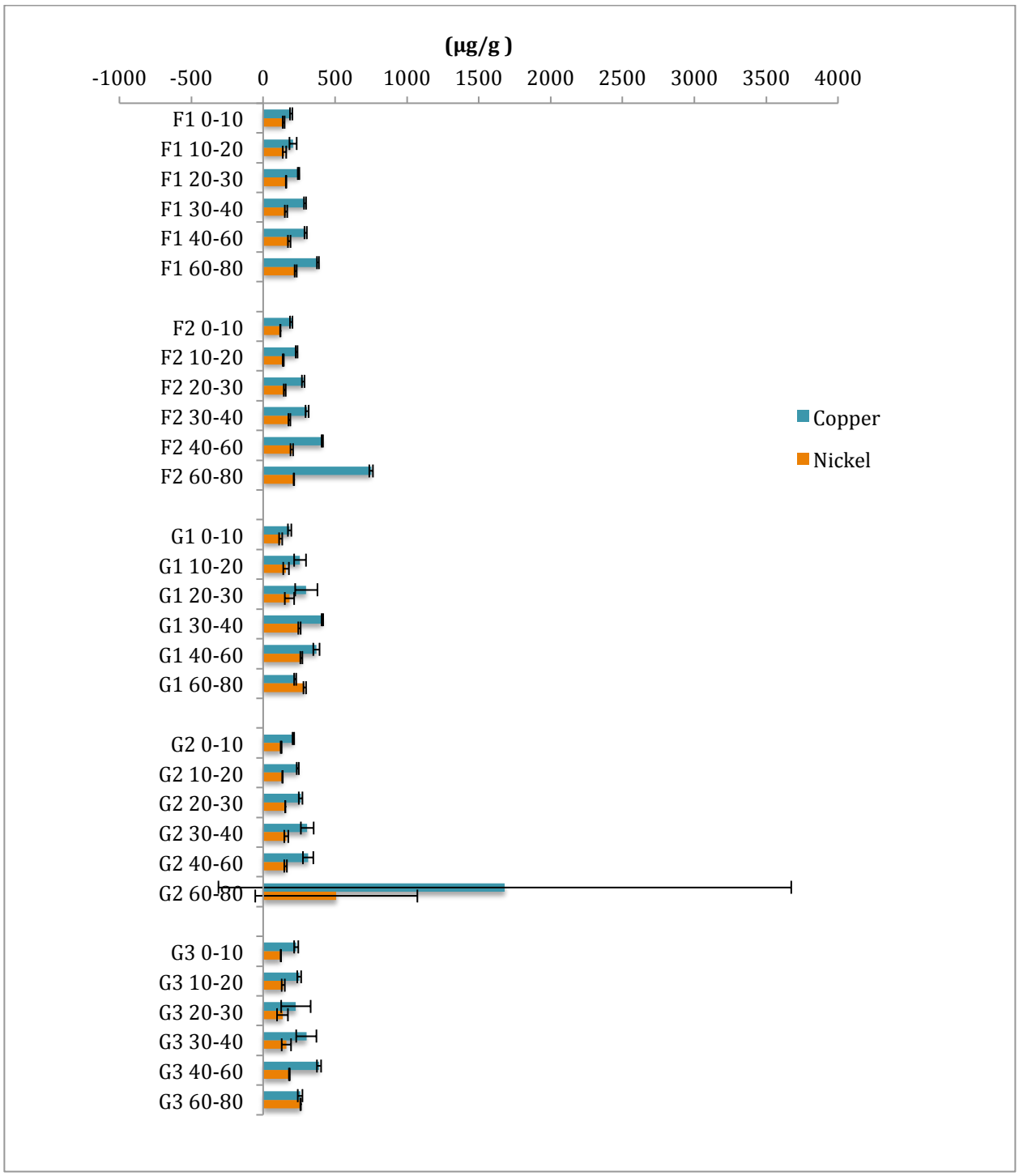


Figure 3.14. Predicted concentrations of nickel and copper in F- and G-transect cores

Correlations between contaminants and toxicity

A Pearson Product Moment Correlation was performed that included concentrations of metals at depth within sediment cores, and results were compared to those of a correlation test performed only on surface samples in Chapter 1. Chromium and nickel were correlated with mortality in surface sediments and within cores. When core segments were assessed, zinc showed a statistical correlation with mortality, which was not revealed by the surface analysis. All correlations were improved by including samples at depth, and results indicate that although zinc may not be a driver of toxicity in surface sediments, it may be significantly correlated with toxicity at depth.

Table 3.3. Comparison of mortality-metal correlations between surface and core samples

Element	**Surface**	**Cores**
Zinc	0.10706 (p>0.05)	**0.42776 (p<0.05)**
Chromium	**0.50568 (p<0.05)**	**0.59844 (p<0.05)**
Copper	0.08299 (p>0.05)	0.05674 (p>0.05)
Nickel	**0.4653 (p<0.05**	**0.53383 (p<0.05)**

DISCUSSION

An exploration of various sample preparation techniques showed comparability between loose powder and packed powder samples. Pellets showed substantially higher readings when they were analyzed without the Prolene® sample cup; however, when measured in the sample cup, they were within the range of the measurements for the two powdered forms. The higher concentrations observed without the sample cup may be due to decreased distance between the x-ray beam and the sample, or due to the fact that Prolene® sample cups interfere with the signal (Drake, 2016). Wet samples were slightly

lower than their dry counterparts, likely due to their difference in composition (more water leads to diluted metal concentrations). However, the differences were not extreme, and indicate that the use of the XRF device in the field unprepared samples may still yield useful information on relative contaminant loads.

Difficulties were encountered throughout the XRF calibration process. The instrument is pre-loaded with several calibration spectra created by the company, with built-in corrections for inter-element interactions. However, the Mudrock calibration curve – typically applicable to sediments from estuarine environments – did not have standards sufficiently high in zinc and chromium to accommodate Bear Creek samples. By supplementing the built-in calibration curve with spiked metals standards, correlations were obtained between XRF-measured results and ICP-MS measured results for zinc, chromium, copper and nickel. For several samples, standard deviations were large, indicating that further factors need to be corrected for within our set of samples. Further analysis is required to determine if other corrections can be applied to improve these correlations. As a first step, a smaller segment of samples from the field will be spiked with metal concentrations in order to assess recovery and determine exact concentrations for the creation of a calibration curve. Though the previous attempt at calibration with in-system samples was ineffective, it is possible that further refinement of this method will yield more accurate results due to the closer matching of calibration standards to sample material.

Using the XRF calibration curves created from in-system samples, metals concentrations in core segments were calculated. As the surface concentration prediction experiment demonstrates, exact quantification of metals by XRF requires improvements

in calibration curves. However, relative concentrations of metals within various cores can be assessed. In general, higher concentrations of metals were found in the most toxic core, C1. Each metal had a unique depth profile. In most cases, zinc and chromium were higher at depth than within surface segments.

Two Pearson Product Moment Correlations were conducted on various subsets of samples. ICP-MS derived metal concentrations and mortality correlations were assessed, and the test determined that mortality was positively correlated with both chromium and nickel to a statistically significant degree. Mortality was negatively correlated with lead and cadmium and positively with arsenic and cadmium, though these relationships were not statistically significant. Correlations were also strong between the key metals within the system, indicating that they are co-varying within the system. In the second correlation test, core samples assessed with handheld XRF compared with mortality values. Mortality was once again significantly associated with chromium, but also with zinc, which had previously not shown a significant relationship. Results indicate that chromium is most strongly correlated with toxicity, though a correlation between zinc and toxicity emerges at depth. The relationship between mortality and chromium is strengthened by the inclusion of core segments, and a correlation between zinc and mortality emerges when sediments at depth are analyzed.

Previous mechanistic studies have assessed chromium at the site, and determined that its oxidation state (Cr(III)) is not conducive to bioavailability or resulting toxicity (Graham et al., 2009) This reveals one limitation of the XRF method – it does not allow for the assessment of metal speciation, which is highly relevant to determining toxicity. Though chromium may not be the driver of toxicity at the site, it is the strongest correlate

with toxicity that is likely to be caused by multiple stressors at the site. However, an analysis of hydrogen sulfide concentrations at the surface and at depth, perhaps conducted without introducing oxygen to the system, could shed light on the bioavailability and toxicity of the metals at the site. Additionally, though high sulfide levels would likely bind metals, hydrogen sulfide itself could be a contributor to sediment toxicity.

Results from this investigation demonstrate the utility of handheld XRF as a fast and field-portable tool for analysis of contaminated sediments. Complementary analyses of metals in Bear Creek sediments by ICP-MS and handheld XRF yielded reasonable correlations, with the handheld XRF technique offering several advantages over ICP-MS. Handheld XRF requires minimal sample preparation, and the rapid analysis of dozens of samples. The handheld XRF technique enables the rapid detection of sediment contaminants and provides a means for pollution mapping in real time and in the field, and at a substantially reduced cost when compared with traditional sediment analysis methods.

Limitations still exist for this method. For example, future work must be completed to find corrections that will improve correlations between ICP-MS and XRF at Bear Creek. For this reason, handheld XRF data remains semi-quantitative in this study, limiting its ability to be used within a regulatory context. Future work will aim to reduce the limitations of this method and bolster the quality of handheld XRF data so it may play a more substantial role in environmental management. Regardless, handheld XRF provides a means to quickly screen sediments for contaminants and pinpoint areas for more in-depth analysis. The use of handheld XRF in this study enabled the assessment of

a much higher number of samples than could be assessed with ICP-MS, including the assessment of metals concentration at depth within sediment cores. Though cause and effect relationships between contaminants and toxicity remain elusive, handheld XRF provided valuable insight into the depth profile of contamination within Bear Creek.

Discussion, Conclusions and Future Directions

Results from the spatial assessment of surface toxicity largely confirmed the first hypothesis of the study: that sediment toxicity would decrease with distance from the shore of the Sparrows Point industrial complex. An assessment of nine sediment cores to depths of 80 cm provided a depth profile for toxicity at the site, but did not support the initial hypothesis that toxicity would uniformly increase with depth. In general, the data support a conclusion that toxicity at the surface cannot be used to adequately predict toxicity at depth, and that sediment toxicity tests run both on surface sediments and at depth are a valuable means to delineate the boundaries of impacted sediments within the Bear Creek system.

An analysis of chemical constituents and their relationships with toxicity leaves one overwhelming conclusion: that no single constituent can be singled out as a likely driver of toxicity, as contaminants tend to vary along with one another. Establishing cause and effect relationships between contaminant classes and toxicity is difficult in complex environmental samples where multiple contaminants may be to blame. In such a

scenario, toxicity tests are especially valuable as a measure of direct biological impacts, because our ability to extrapolate biological impacts from chemical parameters is limited if cause and effect relationships have not been established.

While no single contaminant is likely to blame, several were elevated above screening levels throughout the system. Zinc and chromium uniformly exceeded predicted effects concentrations throughout the system, while copper, cadmium and nickel were all in excess of the PEC values in more than 50% of sites tested. Total sediment PAHs, along with PCBs and TPH, showed exceedances within the system. A Pearson Product Moment Correlation was performed to assess relationships between mortality and contaminant concentrations. Of the constituents assessed, mortality showed a significant positive correlation with nickel, chromium, PCBs, and TPH. In assessing the spatial distribution of contaminants and predicted contaminant concentrations through kriging, the epicenter of pollution was placed at the outlet of Tin Mill Canal, a known point of discharge (EA, 2011). Interpolation also demonstrated an estimated hot spot of metals contamination within Peach Orchard Cove, where minimal toxicity was demonstrated, likely due to their lack of bioavailability. Though this area is not known to be toxic to benthic organisms, high concentrations of metals in the cove indicate a potential need for future analysis of risks to ecological and human health.

Throughout this study, the value of two fast, low-cost, and potentially field-portable analytical chemistry techniques was demonstrated. Excellent correlations were reported between the antibody-based biosensor and GC-MS measurement of PAHs. While handheld XRF correlations left room for improvement, the technique still shows promise for pollution mapping in Bear Creek, and will continue to be developed. In both

cases, the use of these fast methods with simple sample preparation requirements enabled the testing of a greater number of samples than would have been possible with the traditional chemistry techniques to which they were compared. While toxicity at the site is a complex issue, incorporating subsurface samples lead to improved correlations between PAHs, chromium, nickel and sediment mortality.

While both techniques proved valuable in expanding the chemical data set for Bear Creek sediments, there were limitations to the investigation. In the case of handheld XRF, oxidation states are not able to be determined. Similarly, the instrument is unable to detect elements below sodium, and is therefore unable to factor these elements into its conversion of signal responses to weight/weight concentrations. For sediments like these, with high levels of organic pollutants, this may seriously confound efforts to obtain quantitative results from handheld XRF. Several measured contaminant classes correlate with mortality at the site, but sediment toxicity may also derive from organic constituents that were not measured in this study. Previous studies have commented on the substantial oil and tar content of Bear Creek sediments, which has interfered with accurate quantification of organic contaminants in the past (McGee et al., 1999; Klosterhaus & Baker, 2006). Organic contaminants in non-aqueous phases may be a significant driver of toxicity at the site. Further experiments in a controlled laboratory setting may be most helpful in determining the causative agents of toxicity at the site.

In the introduction to Chapter One, two ongoing risk assessments were introduced – one by EA Engineering, and the other by Exponent. Recent data has been released for both studies, which provides an updated risk context for the data obtained in this study.

Since their interim report in early 2015, which demonstrated continued elevated levels of PAHs, metals and PCBs in the offshore area at Sparrows Point, EA Engineering has proceeded with the second phase of their sampling and analysis. They reported the highest values of nickel and zinc at the outlet of the Tin Mill Canal. At this location (Transect C/D in the present study), they proceeded to collect samples in transects across the channel, which demonstrated elevated levels of metals. They also reported high levels of total PCBs and PAHs in these sediments, along with observing samples with "a sheen and/or odor indicating likely petroleum contamination" (EA Engineering, 2016). As a result of high contamination near the Tin Mill Canal outlet, they focused sediment coring efforts in this area. They found varying trends of contaminant concentrations throughout the cores, with some contaminants increasing and some decreasing at depth, similar to the toxicity results reported in this study. Our study supplements these efforts by providing data for surface sediments and sediment cores within the northern reaches of Bear Creek, where high levels of metals were observed and substantial toxicity was seen at depth.

EA Engineering used the data collected 2014-2015, as well as results from fish and crab tissue collected during the Coke Point risk assessment published in 2011 (EA Engineering). They determined that in the northeast/near shore region of the channel, potential carcinogenic risks exist for modeled PAH concentrations for fish and crab ingestion. No risks were determined for wildlife. In the southern portion of the site, potential human health risks and risks to wildlife are indicated for PAHs and PCBs through ingestion of fish and crab tissue. All human health risks related to fish and crab consumption were based on modeled data. Future collection of data within these regions,

as well as within Peach Orchard Cove in the vicinity of Site G3, will be valuable for improving these risk assessments.

Exponent released a technical memorandum in October 2015, which compared all previously collected sediment and surface water data to screening criteria for human health. They found exceedances of screening values for PAHs in surface water at some locations, though the exceedances were only slightly above the screening level. They identified concentrations of PCBs in sediments that exceeded site-specific screening levels developed for the Coke Point assessment. In terms of metals, they identified arsenic as the most noteworthy metal, with sediment concentrations exceeding the EPA's regional screening levels. However, they determined that arsenic did not exceed site-specific levels, which were based on an estimate of actual use and potential exposure. They found elevated levels of chromium, as all previous studies have identified; however, uncertainties still exist surrounding the oxidation state and resultant toxicity of chromium within the system (Exponent, 2015). In conclusion, they determined that the concentrations of arsenic and chromium, as well as PCBs, may warrant further evaluation of human health risks at the site (Exponent, 2015).

Through this study, a number of scientific uncertainties and directions for future research have been identified, with respect to understanding the cause and extent of toxicity within Bear Creek. However, a particularly rich area for future research is in exploring our gaps in knowledge with respect to human health exposure. From the perspective of human health risk assessment, gaps still remain in our knowledge of recreation and usage at the site. Through risk assessments at Coke Point (EA, 2011) and in Bear Creek (EA, 2015; Exponent, 2015), potential risks to recreational users and

watermen have been identified. However, the risks identified in these studies were based on incomplete knowledge of usage patterns. In most risk assessments, site-specific guidelines have been established that are lower than EPA's regional screening levels, given that the Sparrows Point offshore area is considered to be an 'industrial' environment with comparatively low usage. A comprehensive survey to identify the frequency at which Bear Creek is used for fishing, crabbing and recreation would significantly improve our ability to estimate potential risks from consumption or contact to sediments. This data may be particularly relevant to community stakeholders, who are witnessing the beginning of a redevelopment process at a site whose risks just beginning to be understood.

Bear Creek is a known impacted area with demonstrated toxicity, a complex mixture of organic and inorganic contaminants, and potential risks to ecological and human health. It is also an area facing redevelopment decisions in the near future, a process for which a clear understanding of contamination at the site is critical. Risks to human and ecological health have so far been analyzed with respect to current conditions, and not under any particular redevelopment scenarios, which will need to be considered in a risk assessment context. One major finding from this study is the demonstration of increased toxicity and contaminant levels at depth within sediment cores. Under future development scenarios where sediment may be moved, disturbed or re-suspended, these subsurface reservoirs of historic contamination may contribute to increased risks. Overall, this study improves our understanding of the spatial extent of toxicity and contamination at Bear Creek, supplementing ongoing risk assessments while laying the groundwork for future investigations into potential impacts on ecological and human health.

Appendix A – Sampling Coordinates

Table 1. Sampling Coordinates – Surface – June 3rd-4th, 2015

Site Name	Latitude	Longitude	Water Depth (feet)
A1	39.221805°	-76.500591°	15
B1	39.224409°	-76.500852°	13
B2	39.224226°	-76.504357°	11
C1	39.228077°	-76.493229°	5
C2	39.227752°	-76.497105°	10
C3	39.227459°	-76.500977°	15
C4	39.227083°	-76.504547°	9.5
D1	39.230168°	-76.493975°	8
D2	39.229961°	-76.496959°	11
D3	39.229931°	-76.500837°	9
E1	39.233082°	-76.494894°	10
E2	39.232945°	-76.498168°	13
E3	39.232750°	-76.501257°	7
F1	39.235989°	-76.495340°	11
F2	39.235927°	-76.498700°	12
G1	39.240980°	-76.491990°	11
G2	39.240482°	-76.496783°	12
G3	39.241321°	-76.502537°	11
H1	39.243898°	-76.490956°	13
H2	39.246325°	-76.495016°	11
34	39.25344	-76.47736	9
23	39.22209°	-76.50574°	9
24	39.21835°	-76.50709°	14

Table 2. Sampling Coordinates – Surface Repeat – July 28th 2015

Site	Latitude	Longitude	Water Depth (feet)
C1	39.22795	-76.49341	4
D1	39.23008	-76.49491	6
D2	39.22990	-76.49667	9

Table 3. Sampling Coordinates – Cores – July 28th 2015

Core	Latitude	Longitude	Water Depth (feet)
C1	39.22803	-76.49341	5.5
C2	39.22780	-76.49765	10.5

C3	39.22749	-76.50143	14
C4	39.22721	-76.50425	9.5
F1	39.23639	-76.49577	10.5
G1	39.24111	-76.49205	10.5
Core	**Latitude**	**Longitude**	**Water depth (feet)**
G3	39.24139	-76.59240	8
G2	39.24051	-76.49641	9
F2	39.23600	-76.49843	11

Table 4. Sampling Coordinates – Repeat – January 9th, 2016

Core	**Latitude**	**Longitude**	**Water depth (feet)**
G1	39.24115	-76.49203	10
C1	39.22800	-76.49340	5.25
C3	39.22741	-76.50140	14

Appendix B - Statistical results for Chapter 1 Toxicity Tests

For statistical analysis, survival values were converted to decimal proportions, to which an arcsin square root transformation was applied. The reported standard deviations of replicates are for these transformed data. Samples are compared to both the Bigwood Cove control sample, and the in-system control, and p-values are reported for both significant and insignificant comparisons.

Table 1. Mean percent survival of *Leptocheirus plumulosus* in 10-day sediment toxicity test. Tests sites were compared to two control samples, one from within the Patapsco River system in an un-impacted area (Control (IS)). Statistically significant differences are bolded in the table.

Site	Mean % Survival	Standard Deviation of Replicates	Test Site vs. Control	Test Site vs. Control (IS)
CONTROL	98	2.74	---	No (1.000)
CONTROL (IS)	90	17.32	No (1.000)	---
A1	**34**	12.50	**Yes (0.004)**	**Yes (0.024)**
E1	**8**	8.37	**Yes (<0.001)**	**Yes (<0.001)**
E2	**42**	12.04	**Yes (0.006)**	**Yes (0.041)**
E3	67	14.83	No (0.356)	No (1.000)
F1	**28**	13.04	**Yes (<0.001)**	**Yes (0.005)**
F2	**54**	17.10	**Yes (0.049)**	No (0.234)
G1	69	14.75	**Yes (0.045)**	No (1.000)
G2	67	9.08	No (0.349)	No (1.000)
G3	91	8.22	No (1.000)	No (1.000)

H1	76	16.73	No (1.000)	No (1.000)
H2	83	9.75	No (1.000)	No (1.000)

Of the eleven test sites included in the first toxicity test, five – A1, E1, E2, F1, and F2 - showed significant differences from the Bigwood Cove control sediment. All sediments that differed significantly from the Bigwood Cove sediment were also significantly different from the in-system control sediment, except for sample F2 (54% survival). For this test, sediments with above 60% survival did not show significant differences from either control.

Table 2. Mean percent survival of *Leptocheirus plumulosus* in 10-day sediment toxicity test. Tests sites were compared to two control samples, one from within the Patapsco River system in an un-impacted area (Control (IS)). Statistically significant differences are bolded in the table.

Site	Mean % Survival	Standard Deviation of Replicates	Test Site vs. Control	Test Site vs. Control (IS)
CONTROL	99	2.24	---	NO (1.000)
CONTROL (IS)	98	2.74	NO (1.000)	---
B1	**13**	11.09	**YES (0.032)**	**YES (0.042)**
B2	49	21.62	NO (1.000)	NO (1.000)
C1	**1**	2.24	**YES (<0.001)**	**YES (<0.001)**
C2	**9**	8.94	**YES (0.009)**	**YES (0.012)**
C3	29	17.82	**YES (0.032)**	NO (0.396)
C4	50	14.14	NO (1.000)	NO (1.000)
D1	**0**	0.00	**YES (<0.001)**	**YES (<0.001)**

D2	**2**	2.74	**YES (<0.001)**	**YES (0.001)**
D3	**5**	5.00	**YES (0.003)**	**YES (0.004)**
A2	61	13.87	NO (1.000)	NO (1.000)
A'1	42	29.07	NO (0.966)	NO (1.000)

Sediments were also analyzed within their respective transects (A through H). Three transects – B, G and H – passed normality tests (Shapiro-Wilk) and equality of variance tests, and were able to be analyzed parametrically using a one-way ANOVA with post-hoc Holm-Sidak comparisons. The transect-specific parametric analyses increased the power of the test sufficiently to detect differences from control sediments in sites B2, G1, G2, H1, and H2, which were not detected in the less powerful nonparametric analysis. Site G3 remained indistinguishable from the control sample even with parametric analysis.

Table 3. Comparison of sites in transect B to Test 1 control sample using parametric statistics. Statistically significant differences are bolded in the table.

Site	Mean % Survival	Test Site vs. Control
CONTROL	99	---
B1	**13**	**YES (<0.001)**
B2	**49**	**YES (<0.001)**

Table 4. Comparison of sites in transects G and H to Test 2 control sample using parametric statistics. Statistically significant differences are bolded in the table.

Site	Mean % Survival	Test Site vs. Control
CONTROL	98	---
G1	**69**	**YES (<0.001)**
G2	**67**	**YES (<0.001)**
G3	91	NO (0.076)
H1	**76**	**YES (0.003)**
H2	**83**	**YES (0.007)**

Table 5. Core C2 Mean survival and statistical significance in core segments. Bolded values were statistically different from the control sample, with p values in parentheses.

Site	Mean % Survival	Test Site vs. Control
CONTROL	93	---
0 – 10 cm	8	NO (0.052)
10 – 20 cm	**4**	**YES (0.019)**
20 – 30 cm	**3**	**YES (0.010)**
30 – 40 cm	16	NO (0.325)
40 – 60 cm	14	NO (0.251)
60 – 80 cm	6	NO (0.067)

In Core C2, a band of statistically significant toxicity was observed between 10 and 30 cm. Though toxicity was substantial in other segments, the nonparametric nature of the data required rank-based non-parametric analysis resulting in the detection of only a few significantly toxic segments. The core tests were also run with only four replicates, resulting in a loss of statistical power.

Table 6. Core C3 - Mean survival and statistical significance in core segments. Bolded values were statistically different from the control sample for nonparametric analysis, and italicized values were statistically significant from the control sample for nonparametric analysis, with p values in parentheses.

Site	Mean % Survival	Test Site vs. Control: Nonparametric	Test Site vs. Control: Parametric
CONTROL	93	---	
0 – 10 cm	85	NO (1.000)	NO (0.137)
10 – 20 cm	*51*	*NO (0.076)*	*YES (<0.001)*
20 – 30 cm	***45***	***YES (0.040)***	***YES (<0.001)***
30 – 40 cm	***25***	***YES (<0.001)***	***YES (<0.001)***
40 – 60 cm	*51*	*NO (0.096)*	*YES (<0.001)*
60 – 80 cm	*60*	*NO (0.251)*	*YES (0.001)*

In Core C3, nonparametric analysis detected two segments, 20-30 cm and 30-40 cm, that were statistically different from the control, with survivals of 45% and 25%, respectively. Parametric analysis detected three more statistically toxic segments, with only the top layer indistinguishable from the control sample.

Table 7. Core C4 - Mean survival and statistical significance in core segments. Bolded values were statistically different from the control sample, with p values in parentheses.

Site	Mean % Survival	Test Site vs. Control
CONTROL	93	---
0 – 10 cm	74	NO (0.498)
10 – 20 cm	78	NO (0.066)
20 – 30 cm	76	NO (0.170)
30 – 40 cm	85	NO (1.000)
40 – 60 cm	89	NO (1.000)
60 – 80 cm	94	NO (1.000)

In Core C4, percent survival ranged from 73 to 94 percent. None of the core segments were significantly different from the control. Cores F1, F2, G1, G2 and G3 showed similar statistical results to core C2, with some segments detected as statistically significant and some detected as indistinguishable from the controls. The ranges of percent survival that corresponded with significant toxicity were different for each core, due to the nonparametric analysis of the data. Results are presented in tables 1.10 through 1.14.

Table 8. Core F1 - Mean survival and statistical significance in core segments. Bolded values were statistically different from the control sample, with p values in parentheses.

Site	Mean % Survival	Test Site vs. Control
CONTROL	92	---
0 – 10 cm	6	NO (1.000)
10 – 20 cm	**0**	**YES (0.013)**
20 – 30 cm	**0**	**YES (0.013)**
30 – 40 cm	1	NO (0.059)
40 – 60 cm	3	NO (0.214)
60 – 80 cm	**0**	**YES (0.013)**

Table 9. Core F2 - Mean survival and statistical significance in core segments. Bolded values were statistically different from the control sample, with p values in parentheses.

Site	Mean % Survival	Test Site vs. Control
CONTROL	98	---
0 – 10 cm	86	NO (1.000)
10 – 20 cm	49	NO (0.388)
20 – 30 cm	**33**	**YES (0.021)**
30 – 40 cm	**13**	**YES (<0.001)**
40 – 60 cm	**35**	**YES (0.042)**
60 – 80 cm	**33**	**YES (0.034)**

Table 10. Core G1 - Mean survival and statistical significance in core segments. Bolded values were statistically different from the control sample, with p values in parentheses.

Site	Mean % Survival	Test Site vs. Control
CONTROL	92	---
0 – 10 cm	18	NO (1.000)
10 – 20 cm	3	NO (0.055)
20 – 30 cm	**0**	**YES (0.010)**
30 – 40 cm	**0**	**YES (0.010)**
40 – 60 cm	**1**	**YES (0.040)**
60 – 80 cm	3	NO (1.000)

Table 11. Core G2 - Mean survival and statistical significance in core segments. Bolded values were statistically different from the control sample, with p values in parentheses.

Site	Mean % Survival	Test Site vs. Control
CONTROL	98	---
0 – 10 cm	88	NO (1.000)
10 – 20 cm	60	NO (0.649)
20 – 30 cm	53	NO (0.273)
30 – 40 cm	**36**	**YES (0.017)**
40 – 60 cm	**21**	**YES (<0.001)**
60 – 80 cm	**21**	**YES (<0.001)**

Table 12. Core G3 - Mean survival and statistical significance in core segments. Bolded values were statistically different from the control sample, with p values in parentheses.

Site	Mean % Survival	Test Site vs. Control
CONTROL	98	---
0 – 10 cm	63	NO (0.855)
10 – 20 cm	58	NO (0.661)
20 – 30 cm	**33**	**YES (0.006)**
30 – 40 cm	**25**	**YES (<0.001)**
40 – 60 cm	**29**	**YES (0.002)**
60 – 80 cm	61	NO (0.929)

Appendix C - Supplemental Kriging Results

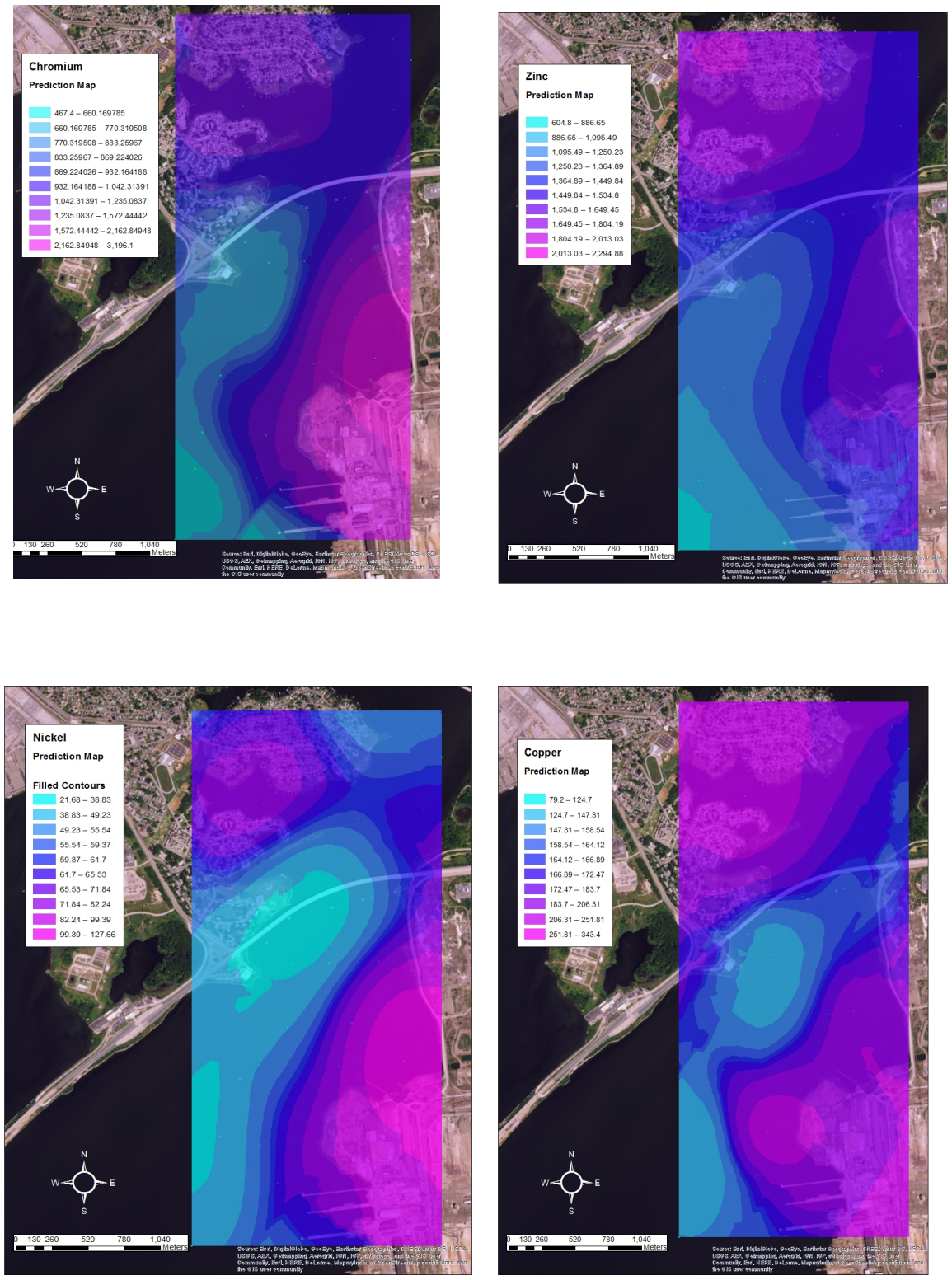


Figure 1. Prediction surfaces: Chromium, zinc, nickel and copper in Bear Creek

Kriging can also be used to estimate and display the probability of exceeding a specified threshold within a system. Probability maps were generated for copper, nickel, zinc and chromium; only copper and nickel are displayed, because chromium and zinc were elevated above the PEC throughout the entire system.

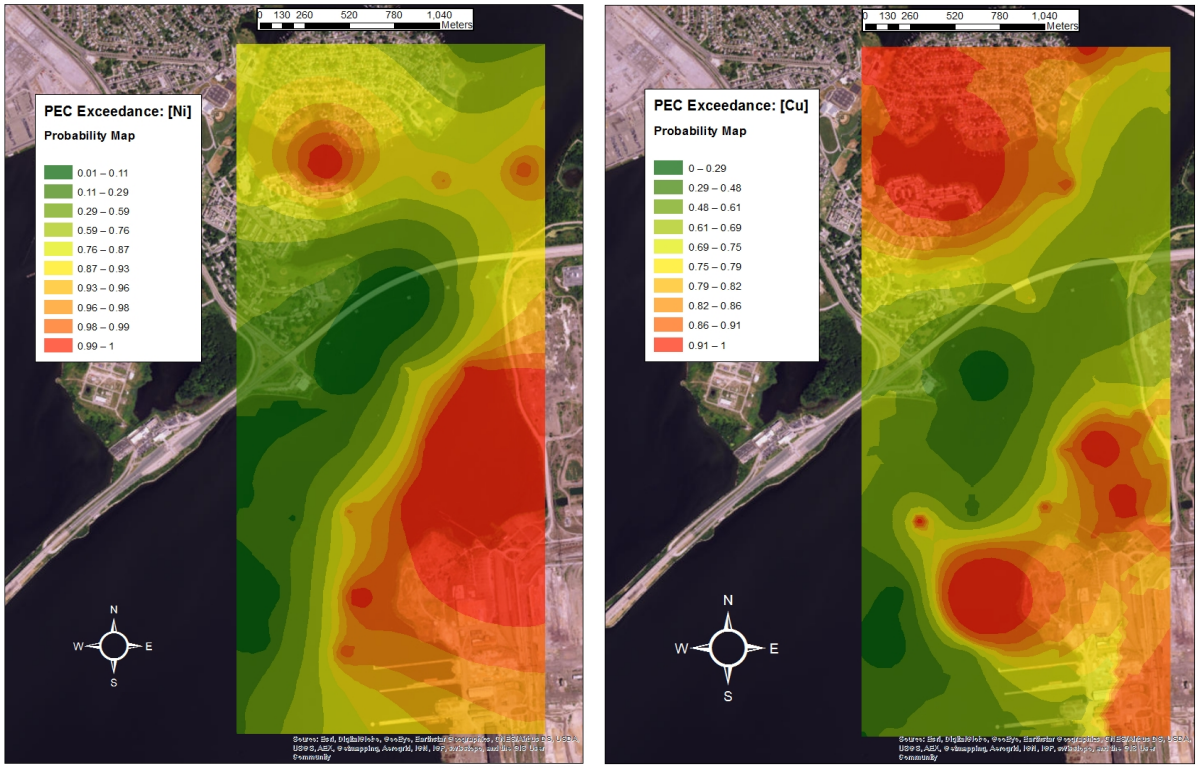


Figure 2. Maps showing probability of exceeding PEC value for nickel and copper

In general, results of spatial interpolation for metals correspond with the toxicity prediction surface, with higher concentrations estimated and PEC exceedances more likely in the near-shore area. However, spatial interpolation demonstrates the potential existence of metals-contaminated hot spots in areas where toxicity was not observed or predicted.

Bibliography

Ashley, J. T.F. and Baker, J. E. (1999), Hydrophobic organic contaminants in surficial sediments of Baltimore Harbor: Inventories and sources. *Environmental Toxicology and Chemistry,* 18: 838– 849. doi: 10.1002/etc.5620180505

ASTM (1992) Standard guide for conducting 10-day static sediment toxicity tests with marine and estuarine amphipods. ASTM Designation E 1367 - 90. *ASTM Annual Book of ASTM Standards, Vol. 11.04*. Amer. Soc. Testing Materials, Philadelphia, PA

ASTM (1994) Standard guide for collection, storage, characterization and manipulation of sediments for toxicological testing. ASTM Designation E 1391 - 90. *ASTM Annual Book of Standards Vol. 11.04* Amer. Soc. Testing Materials, Philadelphia, PA.

Agency for Toxic Substances and Disease Registry (1995). Toxicological Profile for Polycyclic Aromatic Hydrocarbons (PAHs). Department of Health and Human Services, Atlanta, GA, USA

Agency for Toxic Substances and Disease Registry (1999). Toxicological Profile for Total Petroleum Hydrocarbons (TPH). Department of Health and Human Services, Atlanta, GA USA.

Baker, J., Mason, R., Cornwell, J., and Ashley, J., Halka, J. and Hill, J. (1997) Spatial mapping of sedimentary contaminants in the Baltimore Harbor/Patapsco River/Back river system. Solomons, MD, University of Maryland Center for Environmental Science. Chesapeake Biological Laboratory, (*UMCES CBL Reference Series, 97-142*) http://aquaticcommons.org/2883/

Chesapeake Bay Program (1997). Chesapeake Bay Regional Action Plans Development Guidelines. EPA 903-R-97-012. *Chesapeake Bay Program Publications.* Accessed April 20th, 2016 at http://www.chesapeakebay.net/publications/title/chesapeake_bay_regional_action_plans_development_guidelines1

Chesapeake Bay Foundation (2009). Re: Sparrows Point Steel Facility – Notice of Violations and Intent to File Suit under the Consent Decree Between the United States, the State of Maryland, and Bethlehem Steel Corporation; the Resource Conservation and Recovery Act; the Comprehensive Environmental Response, Compensation and Liability Act; The Clean Air Act; The Clean Water Act; the Maryland Erosion and Sediment Control Authority; and the Maryland Environmental Standing Act; Environmental Justice Act. Accessed at http://www.cbf.org/Document.Doc?id=326

Drake, L. (2014). XRF User Guide. Accessed April 2, 2016 at http://www.xrf.guru/

EA Engineering (2016). Phase I Offshore Investigation Report for the Sparrows Point Site. Prepared for Sparrows Point Environmental Trust. Accessed March 31st, 2016 at https://www3.epa.gov/reg3wcmd/ca/md/otherdocs/SparrowsPoint_OffshoreInvestigationReport.pdf

EA Engineering, Science, and Technology, Inc. (2014). Work Plan for Offshore Investigation of the Phase I Area of the Sparrows Point Site. EA Project No. 15131.01.http://www.mde.state.md.us/programs/Land/MarylandBrownfieldVCP/mapping/Documents/Final%20Sparrows%20Point%20Offshore%20Investigation%20Work%20Plan.pdf

EA Engineering (2011). Risk assessment of offshore areas adjacent to the proposed Coke Point dredged material containment facility at Sparrows Point. May 23, 2011. http://mpa.maryland.gov/misc/cokepoint/cp.pdf

Exponent (2015). Health-Based Evaluation of Environmental Data from Sparrows Point Site. [Technical Memorandum]. October, 2015. www.cbf.org/document.doc?id=2415

EnviroAnalytics Group (2015). Tin Mill Canal Sediment Sampling and Analysis Plan. Prepared for Sparrows Point Terminal, LLC and Sparrows Point, LLC. Accessed April 24 at http://www.mde.state.md.us/programs/Land/MarylandBrownfieldVCP/mapping/Documents/TMC%20Sediment%20Sampling%20and%20Analysis%20Plan%20Compiled%20Final%202-19-2015.pdf

Environmental Systems Research Group (ESRI), DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community (2016). World Imagery Basemap. ArcGIS Online basemaps.

ESRI (2011). Interpolation toolset concepts. ArcGIS Resource Center. Accessed April 24th at help.arcgis.com

Exponent. (2011). Assessment of Imminent and Substantial Endangerment to Health and the Environment Due to Contaminant Sources, Transport and Potential Exposure of Human and Environmental Receptors. Exponent Inc. – Draft Report.

Fisher DJ, GP Ziegler and LT Yonkos. 2004. Chronic Toxicity of Sediments from the Baltimore Inner Harbor and Bear Creek to *Leptocheirus plumulosus*. Final Report to the Maryland Department of the Environment. Rep. No. WREC-04-04. University of Maryland, Wye Research and Education Center, Queenstown, MD. 13 pp.

Graham, A.M., Wadhawan, A.R. and Bouwer, E.J. (2009). Chromium Occurrence and speciation in Baltimore Harbor sediments and porewater, Baltimore, Maryland, USA.*Environmental Toxicology and Chemistry,* 28(3): 471-480. Accessed Nov 24, 2014 at http://onlinelibrary.wiley.com/store/10.1897/08-

149.1/asset/5620280304_ftp.pdf?v=1&t=i2wlg4xp&s=137815ab5d32e5dae04f7ff33de7e fbbe917d3ef

Hansen, D.J., Berry, W.J., Boothman, W.S., Pesch, C.E., Mahony, J.D., Di Toro, D.M. , Robson, D.L., Ankley, G.T., Ma, D., and Yan, Q. (1996). Predicting the toxicity of metal-contaminated field sediments using interstitial concentrations of metals and acid-volatile sulfide normalizations. *Environmental Toxicology and Chemistry* 15(12): 2080 – 2094. Doi: 10.1002/etc.5620151204

Hartwell, I.S. and Hameedi, J. (2007) Magnitude and Extent of Contaminated Sediment and Toxicity in Chesapeake Bay. Silver Spring, MD, NOAA/NOS/National Centers for Coastal Ocean Science/Center for Coastal Monitoring and Assessment, (NOAA Technical Memorandum NOS NCCOS, 47) http://aquaticcommons.org/2115/

Hawthorne, S.B., Azzolina, N.A., Neuhauser, E.F, and Kreitinger, J.P. (2007) Predicting bioavailability of sediment polycyclic aromatic hydrocarbons to hyalella Azteca using equilibrium partitioning, supercritical fluid extraction, and porewater concentrations. *Environmental Science and Technology,* 41: 6297-6304. http://pubs.acs.org.proxy-um.researchport.umd.edu/doi/pdf/10.1021/es0702162

Higueras, P., Oyarzun, R., Iraizoz, J.M., Lorenzo, S., Esbri, J.M., Martinez-Coronado, A. (2012). Geochemical surveys for environmental studies in developing countries: Testing a field portable XRF instrument under quasi-realistic conditions. *Journal of Geochemical Exploration,* 113: 3-12.

Ho, K.T., Mickinney, R.A., Kuhn, A., Pelletier, M.C., Burgess, R.M. (1997) Identification of acute toxicants in new Bedford harbor sediments. *Environmental toxicology and chemistry,* 16(3): 551 – 558 10.1002/etc.5620160322

Ingersoll, C.G., MacDonald, D.D., Wang, N., Crane, J.L., Field, L.J., Haverland, P.S., Kemble, N.E., Lindskoog, R.A., Severn, C. Smorong, D.E. (2000) Prediction of sediment toxicity using conseusns-based freshwater sediment quality guidelines http://www.cerc.usgs.gov/pubs/center/pdfdocs/91126.pdf

Inouye L. 2014. Public Review Draft – April 17, 2014 DMMP Issue Paper: Implementation of Revised Freshwater Sediment Screening Values**.** Washington Department of Ecology, for the DMMP and RSET agencies. http://www.nws.usace.army.mil/Portals/27/docs/civilworks/dredging/SMARM%202014/FW%20SL%20presentation.pdf

Kenna, T. C., Nitsche, F. O., Herron, M. M., Mailloux, B. J., Peteet, D., Sritrairat, S., & Baumgarten, J. (2011). Evaluation and calibration of a field portable X-ray fluorescence spectrometer for quantitative analysis of siliciclastic soils and sediments. *Journal of Analytical Atomic Spectrometry*, 26(2), 395-405.

Kirman, D.Z., Sericano, J.L., Wade, T.L, Bianchi, T.S., Marcantonio, F., and Kolker, A.S. (2016). Composition and depth distribution of hydrocarbons in Barataria Bay marsh sediments after the Deepwater Horizon oil spill. *Environmental Pollution 214*: 101-113. doi:10.1016/j.envpol.2016.03.071

Klosterhaus, S. and Baker, J. (2006). Toxicity Identification and Long-Term Contaminant Trends in the Baltimore Harbor – Final Report (DRAFT). University of Maryland Center for Environmental Science. Submitted to Technical & Regulatory Services, Maryland Department of the Environment. Obtained via personal correspondence with the Wye Research and Education Center.

Krivoruchko, K. (2012). Emperical Bayesian Kriging Implemented in ArcGIS Geostatistical Analyst. ArcUser Fall 2012 Edition. Accessed April 24tt, 2016 at http://www.esri.com/news/arcuser/1012/empirical-byesian-kriging.html

Li, X., Kaattari, S. L., Vogelbein, M. A., Vadas, G. G., & Unger, M. A. (2016). A highly sensitive monoclonal antibody based biosensor for quantifying 3–5 ring polycyclic aromatic hydrocarbons (PAHs) in aqueous environmental samples. *Sensing and Bio-Sensing Research*, 7, 115-120.

Long, E.R., Dutch, M., Partridge, V., Weakland, S., and Welch, K. (2013). Revision of sediment quality triad indicators in Puget Sound (Washington, USA): I. a Sediment Chemistry Index and targets for mixtures of toxicants, *Integrated Environmental Assessment and Management*, 9(1): 31 10.1002/ieam.1309

Long, E.R. and Morgan, L.G. (1991). The Potential for Biological Effects of Sediment-Sorbed Contaminants Tested in the National Status and Trends Program. NOAA Technical Memorandum NOS OMA 52. Second printing. Seattle, Washington. http://docs.lib.noaa.gov/noaa_documents/NOS/OMA/TM_NOS_OMA/nos_oma_52.pdf

MacDonald D.D., Ingersoll, C.G., and Berger, T. (2000) Development and evaluation of consensus-based sediment quality guidelines for freshwater ecosystems. *Arch Environ Contam Toxicol* 39:20-31.

Marine Board, Commission on Engineering and Technical Systems; Division on Engineering and Physical Sciences (1997). Contaminated Sediments in Ports and Waterways: Cleanup Strategies and Technologies. Doi: 10.17226/5292

Maryland Department of the Environment (2014). Maryland Fish Consumption Advisories: Statewide Fresh Water, Estuarine and Marine Environments. Accessed March 15, 2016 at http://mde.maryland.gov/programs/Marylander/CitizensInfoCenterHome/Documents/Fish%20Consumption%20Docs/Maryland_Fish_Advisories_2014_Web_bluecatedit.pdf

Mason, RP, Kim, E and Cornwell, J. (2004). Metal Accumulation in Baltimore Harbor: current and past inputs. *Applied Geochemistry*, 19: 1801-1825. Accessed Nov 24, 2014 at http://www.sciencedirect.com/science/article/pii/S0883292704000812

Melquiades, F.L. and Appoloni, C.R. (2004). Application of XRF and field portablt XRF for environmental analysis. *Journal of Radioanalytical and Nuclear Chemistry*, 262(2): 533-541.

McGee, B.L., Fisher, D.J, Wright, D.A., Yonkos, L.T., Ziegler, G.P, Turley, S.D., Farrar, J.D., Moore, D.W., Bridges, T.S. (2004). A field test and comparison of acute and chromic sediment toxicity tests with the estuarine amphipod Leptocheirus plumulosus in Chesapeake Bay, USA. *Environmental Toxicology and Chemistry,* 23(7): 1751-1761 10.1897/03-326

McGee, B. L., Fisher,, D. J., Yonkos, L. T., Ziegler, G. P. and Turley, S. (1999), Assessment of sediment contamination, acute toxicity, and population viability of the estuarine amphipod*Leptocheirus plumulosus* in Baltimore Harbor, Maryland, USA. *Environmental Toxicology and Chemistry*, 18: 2151–2160. doi: 10.1002/etc.5620181006

Morris, J., Saalfield, S., and Barranco, F. (2015). Technical Memorandum on Round 1 Sediment Investigation and Plan for Round 2 Investigation Sparrows Point Phase I Area.

Morgan, R. P., & Sommer, S. E. (1979). Polychiorinated biphenyls in Baltimore Harbor sediments. *Bulletin of environmental contamination and toxicology*, 22(1), 413-419.

Paxton, R. (March 21, 2008). Severstal buys U.S. Sparrows Point mill for $810 mln. Reuters. Accessed March 31st, 2016 at http://www.reuters.com/article/severstal-us-idUSL2115367020080321

Perroy, R.L., Belby, C.S., Mertens, C.J. (2014). Mapping and modeling three dimensional lead contamination in wetland sediments of a former trap-shooting range. Science of the Total Environment, 487: 72-81.

Radu, T., & Diamond, D. (2009). Comparison of soil pollution concentrations determined using AAS and portable XRF techniques. *Journal of Hazardous Materials*, 171(1), 1168-1171.

Reible, D. and Lanczos, T. (Eds.) (2006). Assessment and Remediation of Contaminated Sediments. Proceedings of the NATO Advanced Research Workshop on Assessment and Remediation of Contaminated Sediments Bratislava, Slovak Republic 18–21 May 2005**.** *Nato Science Series: IV: Earth and Environmental Sciences* Volume 73.

Reutter, M. (May 25, 2012). Six reasons why the Sparrows Point steel mill collapsed. The Baltimore Brew. Accessed March 31st, 2016 at https://www.baltimorebrew.com/2012/05/25/six-reasons-why-the-sparrows-point-steel-mill-collapsed/

Shen J, Hong, B., Schugam, Y Zhao, Y and White, J. (2012) Modeling of polychlorinated biphenyls (PCBs) in the Baltimore Harbor. *Ecological Modelling* 242:54-68.

Sherman, N. (June 22, 2016). As Sparrows Point demolition continues, plans for rebuilding begin. The Baltimore Sun. Accessed March 31st, 2016 at http://www.baltimoresun.com/business/bs-bz-sparrows-point-tour-20150622-story.html

Sims, J.G. and Moore, D.W. (1995). Risk of pore water hydrogen sulfide toxicity in dredged material bioassays. US Army Corps of Engineers Waterways Experiment Station. Accessed April 25th 2016 at http://www.dtic.mil/dtic/tr/fulltext/u2/a434428.pdf

Spier, C. R., Vadas, G.G., Kaattari, S.L., and Unger, M.A. (2011). Near real-time, on-site, quantitative analysis of PAHs in the aqueous environment using an antibody-based biosensor. *Environmental Toxicology and Chemistry* 30(7): 1557-1563.

Spier, C.R., Bromage, E.S., Harris, T.M., Unger, M.A. and Kaattari, S.L. (2009). The development and evaluation of monoclonal antibodies for the detection of polycyclic aromatic hydrocarbons. *Anal. Biochem*. 387 (2009), pp. 287–293

Swartz (1999). Consensus Sediment Quality Guidelines for polycyclic aromatic hydrocarbon mixtures. *Environmental Toxicology and Chemistry,* 18(4): 780-787 http://www.waterboards.ca.gov/water_issues/programs/tmdl/docs/303d_policydoc

ter Laak, T.L., Barendregt, A., and Hermens, J.L.M. (2006). Freely dissolved porewater concentrations and sorption coefficients of PAHs in spiked, aged, and field-contaminated soils. Environmental Science & Technology 40(7): 2184 – 2190. DOI: 10.1021/es0524548

Tsai, C. F., Welch, J., Chang, K. Y., Shaeffer, J., & Cronin, L. E. (1979). Bioassay of Baltimore harbor sediments. *Estuaries*, *2*(3), 141-153.

URS Corporation (2010). Baseline Ecological Risk Assessment for On-Site Areas: Draft. Prepared for Severstal Sparrows Point, LLC. Accessed March 31 2016 at http://www.mde.state.md.us/programs/Land/MarylandBrownfieldVCP/ERRP_Superfund/Documents/Sparrows_Point_BERA_Complete_083010.pdf

USEPA (2005). Contaminated Sediment Remediation Guidance for Hazardous Waste Sites. United States Environmental Protection Agency Office of Solid Waste and Emergency Response. Accessed March 31st 2016 at https://clu-in.org/download/contaminantfocus/sediments/contaminated-sediment-remediation-EPA-guidance.pdf

USEPA (2004). The Incidence and Severity of Sediment Contamination in Surface Waters of the United States. National Sediment Quality Survey, Second Edition. EPA-823-R-04-007 http://nepis.epa.gov

USEPA (2002). Clean water action plan: National coastal condition report. EPA 620/R-00-004. U.S. Environmental Protection Agency, Office of Water, Office of Research and Development, Washington, DC. http://nepis.epa.gov

USEPA (2001). Method for Assessing the Chronic Toxicity of Marine and Estuarine Sediment-Associated Contaminants with the Amphipod *Leptocheirus plumulosus.* USEPA Office of Research and Development. EPA/600/R-01/020.

USEPA (1995). QA/QC guidance for sampling and analysis of sediments, water and tissues for dredged material evaluations. EPA 823-B-95-001. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

Wood, P. (September 18, 2014). Sparrows Point owner, government reach clean up agreement. The Baltimore Sun. Accessed March 31st, 2016 at http://www.baltimoresun.com/news/maryland/baltimore-county/dundalk/bs-md-co-sparrows-point-20140918-story.html

Yonkos, L.T. (2015). Draft Report: 2015 Toxicity Testing of Baltimore Harbor Sediments. Prepared for Paul Smail and Dr. Beth McGee, Chesapeake Bay Foundation.

Yonkos, L.T., Ziegler, G.P., and Friedel, E.A. (2012). Toxicity Testing of Baltimore Harbor Sediments. Report # WREC-12-01. Final Report to Chesapeake Bay Foundation, Annapolis MD. 14 pp.

Zhang, Y., Zhang, L., Gao, Z., and Dai, S. (2012). Investigating the quantitative structure-activity relationships for antibody recognition of two immunoassays for polycyclic aromatic hydrocarbons by multiple regression methods. *Sensors*, 12(7): 9363-9374. doi: 10.3390/s120709363