

**ASSESSING THE GEOGRAPHIC DISTRIBUTION OF MERCURY AND LEAD IN  
BUFFALO RIVER SEDIMENTS**

**by**

**Peter S. Rodriguez**

**A major research paper**

**presented to Ryerson University**

**in partial fulfillment of the requirements for the degree of**

**Master of Spatial Analysis (MSA)**

**Toronto, Ontario, Canada**

**© Peter S. Rodriguez 2009**

**Author's Declaration**

I hereby declare that I am the sole author of this major research paper.

I authorize Ryerson University to lend this Research Paper to other institutions or individuals for purposes of scholarly research.

---

Peter S. Rodriguez

I further authorize Ryerson University to reproduce this major research paper by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

---

Peter S. Rodriguez

## **Abstract**

The Buffalo River Area of Concern (AoC) is associated with heavy metal pollution. This paper used sediment core data to model the geographic distribution of mercury (Hg) and lead (Pb) within surface and subsurface sediments of a section of the AoC. Using ordinary kriging, a geostatistical interpolation technique, Hg and Pb concentration prediction maps were created. Standard maps were created by interpolating a prediction surface for the entire study area. In addition, spliced maps were created by interpolating prediction surfaces for three sections of the study area and merging these together to generate a quasi-continuous surface. The latter type of map was needed in order to account for the pronounced meandering nature of the river and to explore the potential existence of regional and local Hg and Pb sedimentation patterns. The results of this research show that there was trend in Hg and Pb contamination that increased from east to west (the direction of river flow), with peak values found in the mid region of the study area. Moreover, surface sediments were less polluted than subsurface sediments for both Hg and Pb. Hg contamination was less widespread but more concentrated compared to that of Pb. Finally, this study highlighted several Hg and Pb contamination hotspots, which could be targeted for future sediment restoration endeavours in the AoC.

## **Acknowledgements**

I would like to acknowledge the support and guidance of my supervisor, Dr. Wayne Forsythe. Also, I want to extend my thanks to Dr. Andrew Millward for his helpful comments and encouragement.

Dr. Kim Irvine and Mary Perrelli from the Department of Geography and Planning, Buffalo State, State University of New York are thanked for providing the original data files which made this research possible.

## **Dedication**

This work is dedicated to my family and friends.

## Contents

Author's Declaration.....	ii
Abstract.....	iii
Acknowledgements.....	iv
Dedication.....	v
List of Tables.....	viii
List of Figures.....	ix
List of Acronyms.....	x
Chapter 1: Introduction.....	1
1.1 Research Objectives.....	3
1.2 Study Area.....	4
Chapter 2: Background.....	8
2.1 Mercury (Hg).....	8
2.2 Lead (Pb).....	9
2.3 River Pollution and Sedimentology.....	11
2.4 TEL and PEL.....	13
2.5 Kriging Analysis.....	14
2.5.1 The Semi-variogram.....	17
2.5.2 Anisotropy.....	21
Chapter 3: Data and Methods.....	22
3.1 Data.....	22
3.2 Exploratory Data Analysis.....	26
3.3 Log-transformation.....	26
3.4 Ordinary Kriging Analysis.....	27
3.5 Model Cross-Validation.....	29
3.6 The Best Model Index (BMI).....	32
3.7 Ordinary Kriging Flow Chart.....	35
Chapter 4: Results.....	38
4.1 Exploratory Data Analysis.....	38
4.2 Ordinary Kriging Analysis.....	44
4.3 Distribution of Hg and Pb in surface sediments.....	47

4.4	Distribution of Hg and Pb in subsurface samples .....	55
Chapter 5:	Discussion .....	60
5.1	Exploratory Data Analysis .....	60
5.2	Semi-variograms and Kriging analysis .....	61
5.3	Standard and Spliced prediction maps .....	66
5.4	Geographical distribution of Hg and Pb in the Buffalo River.....	68
5.5	Study Limitations and Future Work.....	74
Chapter 6:	Conclusions .....	76
References	.....	78
Appendix A	.....	81
Appendix B	.....	82

## List of Tables

<b>Table 2.1.</b> TEL and PEL values for Hg and Pb.....	14
<b>Table 4.1.</b> Summary statistics of surface and subsurface sediment data sets.....	42
<b>Table A.1.</b> The best kriging models and their parameters. ....	81

## List of Figures

<b>Figure 1.1.</b> Study area .....	6
<b>Figure 1.2.</b> Main elements of a meandering river and river section nomenclature.....	7
<b>Figure 2.1.</b> The semi-variogram and its elements.....	18
<b>Figure 3.1.</b> Distribution of sample points within the study area.....	23
<b>Figure 3.2.</b> Distribution of data points within depth levels in the study Area.....	25
<b>Figure 3.3.</b> Scatter plots of prediction error measurements against the best model index (BMI).. .....	34
<b>Figure 3.4.</b> Ordinary kriging methodology flow chart.....	37
<b>Figure 4.1.</b> Average concentration of Hg and Pb in subsurface sediments.....	39
<b>Figure 4.2.</b> Correlation scatter plots of log Pb against log Hg.....	43
<b>Figure 4.3.</b> Prediction error and Q-Q plots.....	47
<b>Figure 4.4.</b> The geographical distribution of Hg in surface sediments in the study area.	49
<b>Figure 4.5.</b> The geographical distribution of Pb in surface sediments in the study area.	50
<b>Figure 4.6.</b> The geographical distribution of Hg in surface sediments in the study area (spliced map).....	53
<b>Figure 4.7.</b> The geographical distribution of Pb in surface sediments in the study area (spliced map).....	54
<b>Figure 4.8.</b> The geographical distribution of Hg in the study area at four depth levels...	58
<b>Figure 4.9.</b> The geographical distribution of Pb in the study area at four depth levels. ..	59
<b>Figure 5.1.</b> The relationship between sill and nugget for all data sets.....	63
<b>Figure 5.2.</b> Prediction map of water depth along the study area.....	71
<b>Figure B.1.</b> Prediction error maps for Hg in the study area at four depth levels.....	82
<b>Figure B.2.</b> Prediction error maps for Pb in the study area at four depth levels.....	83

## List of Acronyms

AoC	Area of Concern
ADBS	Average distance between samples
ASE	Average Standard Error
BMI	Best Model Index
EDA	Exploratory data analysis
EV	Estimated Value(s)
Hg	Mercury
GIS	Geographic Information Systems
M	Mean
MS	Mean Standardized
NM1	First North Meander
NM2	Second North Meander
NM3	Third North Meander
NYSDEC	New York State Department of Environmental Conservation
RMS	Root Mean Square
RMSS	Root Mean Square Standardized
SM1	First South Meander
SM2	Second South Meander
SM3	Third South Meander
SV	Sampled (Measured) Value
Pb	Lead
PEL	Probable Effect Level
PER	Possible Effect Range
TEL	Threshold effect level
USEPA	United States Environmental Protection Agency

## **Chapter 1: Introduction**

Mercury has been used by humans from the very beginning of civilization. Traces of liquid mercury have been found in Egyptian ceremonial cups and the highly prized vermilion pigment, which contains mercury, was used by Romans to decorate their villas (D'Itri and D'Itri, 1977). More recently, mercury has been used in pesticides, paints, ointments and even medical treatments (D'Itri and D'Itri, 1977).

The low melting point, durability and malleability of lead allowed early human societies to use it for construction, cosmetic and even gastronomical purposes. The Egyptians used lead in eye paints (Smol, 2008), whereas the Romans used it to build pipes to supply their cities with water (Harrison and Laxen, 1981) and even used it as a wine sweetener and preserver (Time, 1966). This close relationship between the Romans and lead has prompted researchers to believe that lead poisoning was one of the reasons for the demise of the Roman Empire (Lewis, 1985).

Even though some of the toxic properties of mercury and lead were known by the Romans, it was only in the mid 20<sup>th</sup> century that these adverse health effects became well known to the general public (Botkin and Keller, 2005). Several public health tragedies and recent discoveries have contributed to this awareness. For instance, the environmental pollution of mercury was largely ignored until the onset of the infamous Minamata Bay (Japan) tragedy in the 1950's, when dozens of fishermen died and hundreds became permanently ill after eating mercury contaminated fish that were caught in the bay. A prominent Japanese chemical factory was later found guilty of dumping mercury effluents into the bay (D'Itri and D'Itri, 1977). Similarly, mercury pollution in the Wabigoon-English River System (Canada) caused many native Indians to become

severely ill and prompted widespread public condemnation in the 1960's (D'Itri and D'Itri, 1977). Another infamous tragedy occurred in the early 1970's, when more than 400 people died in Iraq after eating bread prepared with wheat seeds tainted with mercurial fungicides (Smol, 2008). In terms of lead poisoning, it is now known that exposure to this metal lowers the IQ of children (Harrison and Laxen, 1981) and may contribute to increased delinquent behaviour (Botkin and Keller, 2005).

The concentration of mercury and lead in the environment at global and local scales can be gauged by analyzing air, water, soil, sediment and even glacial ice cores. There is debate over whether sediments can provide faithful information with regard to flux in pollutants, and especially that of mercury (Smol, 2008). This debate arises from the fact that natural chemical and physical processes can increase mercury concentration within sediments without the need of new additions into the system (Miller and Orbock Miller, 2007). A similar situation occurs with lead as its depositional patterns and concentration levels are also the result of a myriad of natural processes within water systems (Harrison and Laxen, 1981).

Less controversial than the use of sediments, is the importance of accurately mapping pollutant distribution so that effective preventive and remedial action can be carried out. Riverbed sediments have been shown to work well as indicators of the spatial variation of contaminant levels (Miller and Orbock Miller, 2007). However, the accurate mapping of contaminant levels presents special challenges. For instance, river flow direction, water flow speed, physical pattern of the river and direct and indirect sources of contamination are just few of the factors that influence sediment depositional patterns.

The Buffalo River was contaminated with toxic elements as a result of decades of industrial activity in the surrounding land (Sutton, 2006). The river undergoes periodic dredging which alters natural sediment depositional patterns. A stretch of this river, known as the Buffalo River Area of Concern (AoC) is of special interest due to its high number of pollutants and pollution levels. A sediment survey conducted by the New York State Department of Environmental Conservation (NYSDEC) in 2005, analyzed more than one-hundred different pollutants. The above factors, when combined, provide an ideal study site to carry out the geospatial analysis of pollutants.

Ordinary kriging is a well tested method used to investigate the geographic distribution of pollutants in sediments (Ouyang *et al.*, 2002; Ouyang *et al.* 2003a; Ouyang *et al.* 2003b; Forsythe *et al.*, 2004; Forsythe and Marvin, 2009). This geostatistical method produces interpolation surfaces that are unbiased and exact (Clark, 1979; Wackernagel, 2003). In practical terms, unbiased means that the residuals' average is very close to zero, whereas exact means that estimated values are very close to the sampled values. Unlike other interpolation techniques, ordinary kriging provides a prediction error surface which allows for sound assessment of predicted values. This paper presents ordinary kriging analysis of two pollutants found in the AoC, mercury (Hg) and lead (Pb).

### **1.1 Research Objectives**

The main objective of this paper was to assess and compare the spatial distribution of Hg and Pb concentration in surface and subsurface sediments in a section of the AoC. In

order to achieve this main objective, several intermediate ones first need to be achieved.

These are as follows,

1. Conduct exploratory data analysis (EDA) of Hg and Pb present in the sediment samples.
2. Conduct full horizontal variographic analysis of Hg and Pb present in the sediment samples.
3. Determine if the best way to interpolate Hg and Pb concentration values along a meandering river is to use the entire length or divide it into regional sections. The former approach requires the development of a global kriging model whereas the latter requires several regional kriging models for each pollutant.
4. Determine the location and number of areas within the study area that have pollution levels below the Threshold Effect Level (TEL), between TEL and the Probable Effect Level (PEL), and above the PEL.
5. Produce prediction surfaces and associated standard error maps of the distribution of Hg and Pb concentrations within surface and subsurface sediments.

## **1.2 Study Area**

The Buffalo River runs through to the City of Buffalo which is located in New York State (USA), near the Canadian border (Figure 1.1A). This river flows east to west and discharges into Lake Erie. The land use in the upper basin of the river is mainly agricultural and woodland; whereas the land use in the lower basin is industrial and urban (Shreeram, 2004). Stream bank erosion especially due to ice scouring, is an important

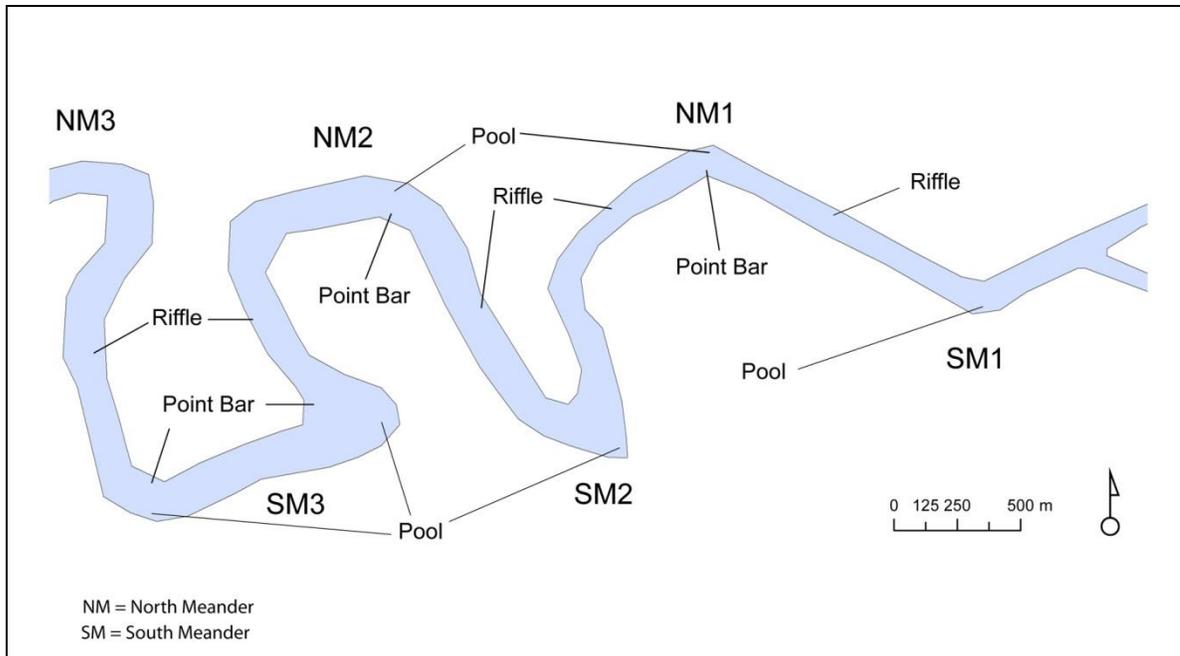
contributor to sediment build-up in the Buffalo River, which receives about 86,000 tons of sediment each year from its three watersheds (Shreeram, 2004).

This paper used data taken from a section of the AoC (yellow ellipse in Figure 1.1B). The AoC is characterized by its heavy industrial pollution, which is mostly found in contaminated sediments as the industries that were present in the AoC for decades are now largely gone (Sutton, 2006). These polluted sediments, both surface and subsurface, together with runoff and atmospheric deposition account for most of the sources of new contamination (US EPA, 2008). Within the vicinity of the AoC and throughout the Buffalo River watershed there are more than 45 inactive hazardous waste sites, 33 combined sewer overflow outfalls and several sewage systems. All of these factors also contribute to the on-going pollution of the AoC (US EPA, 2008). The AoC has a long history of periodic dredging (US EPA, 2008). In fact, in 2003 the AoC was dredged for the last time before sediment samples were collected in 2005 (Sutton, 2006). It is likely that dredging alters the natural sedimentation pattern within the AoC.

As seen in Figure 1.1B the segment of the AoC studied in this paper has a pronounced meandering pattern. In general terms, a meandering river is characterized by having alternate accumulations of sediments called point bars which are opposite to deeper areas called pools, which are in turn separated by shallower and longer areas called riffles (Miller and Orbock Miller, 2007). These river morphological elements together with the river section nomenclature used in this paper are shown in Figure 1.2.



**Figure 1.1.** Study area. The locations of the Buffalo River and the AoC within North America are seen in A. In B a close-up of the AoC is shown in orange colour. The section of the AoC studied in this paper is highlighted with the yellow ellipse. Sources: Figure 1.1A uses a satellite image taken from Google Maps together with ESRI ArcCanada 3.0 (2003) shapefiles. Figure 1.1B was modified from USEPA (2008)



**Figure 1.2.** Main elements of a meandering river and river section nomenclature. The order of the numeration goes from east to west. For instance, SM1 means first south meander.

Meandering rivers have, in addition to their main flow, a second circular flow which is found in the meander itself and is oriented oblique to the main flow (Miller and Orbock Miller, 2007). This means that different pollutant depositional patterns could be expected in the meanders compared to those found in the straighter sections of the river.

Although the study area is located in an urban location and it has been altered by human activity, features such as pools, point bars and riffles are found within it. The reason for this is that these features are the product of the interaction between water flow, sediment transportation and bedforms.

## Chapter 2: Background

### 2.1 Mercury (Hg)

Mercury is naturally found in the environment where it cycles between water, soil and air as the result of biological and chemical reactions (Botkin and Keller, 2005). However, anthropogenic emissions of Hg have added great amounts of this heavy metal to the environment.

Many scientists believe that mercury pollution now threatens the health of people all around the globe (University of Wisconsin-Madison, 2006). For instance, a possible link between increased Hg pollution and increased cases of Autism has been hypothesized in the scientific literature (Forsyth, 2005). Researchers in Texas have found that for every thousand pounds of mercury released to the atmosphere a 61% increase in the rate of Autism can be expected (Palmer *et al.*, 2006). Similarly, Hg affects the immune system as well as increases the risk of cardio vascular diseases in humans (University of Wisconsin-Madison, 2006).

Mercury bioaccumulation starts with the conversion of Hg into methyl-mercury usually by aquatic organisms. Methyl-mercury is then absorbed by fish tissue and finally passes to top predators where it accumulates (Botkin and Keller, 2005). It was this phenomenon that caused Minamata Bay fisherman to have up to 30,000 times more methyl-mercury in their bodies than the water systems in which they fished (D'Itri and D'Itri, 1977).

Only a relative short period of ingestion of Hg contaminated products is needed to cause severe health issues in humans. For instance, if a person eats contaminated fish

everyday for three months, it is likely that this person will lose sight, become crippled or even die (D'Itri and D'Itri, 1977).

Hg has an affinity for fine-grained organic sediments and thus, usually does not move too far away from its point source (Smol, 2008). Hg is found in an elemental and organic form, known as methyl-mercury. The latter form is of greater environmental concern due to its high mobility and propensity to bioaccumulate in living organisms (Smol, 2008).

## **2.2 Lead (Pb)**

Lead enters water systems mainly via runoff, groundwater inflows and atmospheric deposition (Smol, 2008). Once Pb enters a water system it quickly becomes part of the suspended particles and sediments and most of it eventually makes its way into marine sediments (Harrison and Laxen, 1981).

Pb makes its way into the human body through water, air and food. For instance, Pb is mostly released to the atmosphere by smelting industries and until recently by vehicles through the combustion of leaded-gasoline (Harrison and Laxen, 1981). After reaching a peak in 1972, Pb atmospheric emissions in the USA have been reduced by 98% as a direct effect of the Clean Air Act of 1970, which required the removal of lead additives from gasoline (Smol, 2008). Pb is also present in the tap water of homes that have lead piping (Harrison and Laxen, 1981) and even in food at varying concentrations. A study conducted in Spain revealed that lead concentrations in chicken, pork and beef were 0.0069, 0.005 and 0.0019  $\mu\text{g/g}$  respectively (Gonzalez-Weller *et al.*, 2006). Similarly, a study conducted in Hong Kong found that lead concentration in large shrimp

can be as high as 6  $\mu\text{g/g}$  (Cheung and Wong, 2006). Although unlikely in practice, in theory, five kilos of chicken or six grams of shrimp could have enough Pb to cause adverse health effects as concentrations above 35  $\mu\text{g/g}$  lie in the possible effect range (Canadian Council of Ministers of the Environment, 2001)

Pb poisoning can lead to severe damage of the human central nervous system and if the exposure is prolonged, it can even result in death (Harrison and Laxen, 1981). Once Pb enters the body it attaches to red-blood cells, accumulating in soft tissues and in bone (Harrison and Laxen, 1981). Recent studies have shown that antisocial, aggressive and delinquent behaviour in children is associated with above average concentrations of lead in their bones (Botkin and Keller, 2005).

Body burden is the amount of toxic chemicals present in an individual at a given point in time (Botkin and Keller, 2005). The body burden for Hg and Pb in an average human body (70 kg) is 13mg and 150 mg respectively, which reveals the heavier use or exposure of Pb compared to Hg in humans (Botkin and Keller, 2005).

The distribution of Pb in the environment and its toxic effects on living organisms are closely related to its physical and chemical forms (Harrison and Laxen, 1981). For instance, the formation of Pb and organic matter complexes is dependent on pH, as the likelihood of Pb precipitating out of water increases between a pH of 7.5 and 8.5 (Harrison and Laxen, 1981). Hence, at pH values other than between 7.5 and 8.5, Pb concentrations could be relatively low in sediments but relatively high in the water above.

Another spatial characteristic of Pb is that it tends to accumulate in deep clayish sediments in ocean waters showing concentrations between 50 and 150 mg/kg and shows a decrease in concentration towards the coast (Harrison and Laxen, 1981).

### **2.3 River Pollution and Sedimentology**

Only about 3% of the earth's water is fresh (non-saline) and most of this is locked in glaciers or in geological strata as groundwater (Smol, 2008). Similarly, rivers only hold about 0.003% of the earth's available freshwater (Miller and Orbock Miller, 2007). The importance of rivers for the development of human societies has been essential as they have been used as transportation networks and for crop irrigation among other important uses.

Most river pollutants can be classified into either inorganic or organic contaminants. The former include pollutants such as metals, metalloids, nutrients and radionuclides, whereas the latter includes pathogens, microorganisms, pesticides (i.e. dichloro diphenyl trichloroethane (DDT)) and polychlorinated biphenyls (PCBs) among others (Miller and Orbock Miller, 2007).

Metals can be introduced to water systems through several ways including air deposition, rock weathering and runoff from industrial sources (Smol, 2008). Once in the water, metals tend to bind readily with organic and non-organic particles and are deposited in sediments (Smol, 2008). As a result, sediments tend to have higher metal concentrations than the overlying water (Smol, 2008).

Sediments can be used to investigate the health of rivers as well as pollution sources and transport dynamics within river basins (Miller and Orbock Miller, 2007). This use is partly due to the fact that riverbed sediments show less variation in pollutant concentration over time and space than the overlying water (Miller and Orbock Miller, 2007). However, a complication arises as sediments do not remain static over time;

rather they move around a water system by means of chemical/physical and artificial mechanical processes such as ice scouring and dredging.

The propensity of sediments to retain metal contaminants increases with smaller grain size composition and bigger surface area per unit mass (Miller and Orbock Miller, 2007). For a given amount of sediment, total surface area increases with decreasing particle grain size. This means that smaller particles such as clay can hold more metal molecules than bigger particles such as sand or silt (Miller and Orbock Miller, 2007).

Sediment distribution can also be influenced by natural physical processes. For example, water flow speed and particle size interact to produce a pattern of decreasing grain size with decreasing distance from the river bank (Miller and Orbock Miller, 2007). This phenomenon combined with the affinity of smaller particles for metals, could result in higher metal pollution in sediments near riverbanks compared to those found in the middle of the riverbed.

Another depositional pattern is that surface sediments tend to be more polluted than subsurface sediments. For instance, the Cedar-Ortega River system has been shown to have greater concentration levels of Hg (Ouyang *et al.*, 2003a) and DDT (Ouyang *et al.*, 2003b) in surface sediments (0-10 cm) than in deeper ones. Furthermore, other studies indicate that rivers that have had a prolonged influx of Pb, can store (in their sediments) about 50% of the total Pb found within the river's watershed (Bookstrom *et al.*, 2001). This suggests that the length of pollutant exposure can influence the spatial distribution of pollutants within a river's sediments.

Sediment distribution and pollution is also influenced by natural erosion, scouring and filling events (Miller and Orbock Miller, 2007). In the case of the study area,

artificial mechanical disturbance also influences the river's sedimentology as periodic dredging takes place in this section of the Buffalo River.

In summary, the distribution of Hg and Pb within a river is likely to be non-random as sedimentation patterns are the product of specific and well studied processes.

## **2.4 TEL and PEL**

This paper uses the threshold effect level (TEL) and probable effect level (PEL) given by the Canadian Council of Ministers of the Environment (CCME) as indicators of sediment quality in freshwater sediments. The TEL indicates a concentration below which adverse biological effects are expected to occur rarely ( $< 25\%$ ), while the PEL indicates a concentration above which adverse effects are expected to occur frequently ( $> 50\%$ ) (Canadian Council of Ministers of the Environment, 2001). The range between TEL and PEL is known as the possible effect range (PER) and biological abnormalities are expected to occur occasionally within this range.

It should be noted that most pollutants do not yet have official Canadian TEL values; instead these pollutants have an interim sediment quality guideline (ISQG). The ISQG acts as a temporary value until the full battery of toxicology tests are carried out and a TEL designation is published (Canadian Council of Ministers of the Environment, 2001). This paper used the ISQG values as a proxy for the TEL values of Hg and Pb. Table 1 shows the CCME values together with those given by the United States Environmental Protection Agency (USEPA) for comparison. As seen in this table the US toxicology standards are less stringent than the Canadian ones.

**Table 2.1.** TEL and PEL values for Hg and Pb. Concentrations are given in  $\mu\text{g/g}$  dry weight. *Source: CCME (2001) and USEPA (2008).*

Source	Hg		Pb	
	TEL (ISQG)*	PEL	TEL (ISQG)*	PEL
CCME	0.170	0.486	35.0	91.3
USA EPA+	0.230	0.870	47.82	161.06

\* ISQG values are used as proxy for TEL values

+ These values are logistic regression model values and were originally given in ppm

## 2.5 Kriging Analysis

The distribution of environmental phenomena can be mapped through several methods including inverse distance weighting, polynomial interpolation and kriging. The latter has been used to map stream temperature in rivers (Gardner *et al.*, 2003), estuary water quality (Little *et al.*, 1997) and distribution of pollutants in river sediments (Ouyang *et al.*, 2002; Ouyang *et al.*, 2003a; Ouyang *et al.*, 2003b) and lake sediments (Forsythe *et al.*, 2004; Forsythe and Marvin, 2009).

Ouyang *et al.* (2003a) used kriging to map the distribution of Hg in the Cedar-Ortega water system sediments and Ouyang *et al.* (2003b) did the same for DDT. Extensive contamination of Hg was found in the top metre of sediment, whereas DDT was found to be widely distributed in the top half metre of sediment. Similarly, Forsythe *et al.* (2004) used kriging to show that the levels of Hg and Pb in sediments within Lake Ontario and Erie have lowered over a period of three decades. Forsythe and Marvin (2009) used kriging to analyze the sediments of Lake Huron and found a significant reduction in Hg and a minimal reduction in Pb levels over a period of about 30 years.

There are many types of kriging each designed for a specific purpose. For instance, there are simple, ordinary, universal, probability, indicator, and disjunctive

kriging just to name a few. Also, for all the types mentioned, there is the potential for co-kriging equivalent analyses. Co-kriging is a technique that uses two variables, measured at the same or at nearby locations, to create prediction surfaces rather than just one as is the case with kriging. Details on the assumptions and data requirements of each type of kriging can be found in the works by and Johnston et al. (2001) and Wackernagel (2003). As this paper deals with ordinary kriging, only this technique is explained in more detail below.

In simple terms, kriging can be defined as the process of determining the best estimation of an unknown value at some location within a given area of interest (Clark, 1979). Kriging originated from the work done by Daniel Krige in South Africa in the 1950's as part of his master's thesis (Clark, 1979). Krige's work was later included in the Theory of Regionalized Variables developed by French mathematician, Georges Matheron (Clark, 1979).

In classical statistics, observations are assumed to be independent; however, correlation among observations is a defining feature of kriging (Johnston *et al.*, 2001). Another, defining characteristics of kriging is that it uses both mathematical and statistical models. As the exact value at some unknown location can never be known with 100% accuracy, a probability is associated with each predicted value. The statistical model used in kriging allows for the calculation of these probabilities.

At its core kriging uses a simple formula,

$$Z(s) = \mu(s) + \epsilon(s) \quad (1)$$

Where  $\mu(s)$  is a deterministic trend,  $\epsilon(s)$  is the associated error and both terms are function of location (i.e. spatial coordinates). Kriging assumes that the error component

is random and dependent on the displacement between points rather than on their actual locations (Johnston *et al.*, 2001). The former assumption means that the average of the errors of predicted values should be close to zero, whereas the latter assumption means that these errors only depend on the distances between the geographic locations of values (Johnston *et al.*, 2001). The error is assumed to have a specific distribution whereas the fixed trend component can vary depending on the location within the area of interest (Clark, 1979). If the trend varies with location, then universal kriging should be used; however, if the trend is unknown and assumed to be constant then ordinary kriging should be used (Johnston *et al.*, 2001).

Kriging is made up of two processes, the quantification of the spatial structure of the data and the production of the estimated values. Ordinary kriging produces the best linear unbiased estimator whenever there is no trend in the data [i.e. constant  $\mu(s)$ ] and a semi-variogram has been calculated with the proper parameters (Clark, 1979). The process of fitting a semi-variogram to quantify the spatial structure (i.e., spatial autocorrelation) of data is known as variography (Johnston *et al.*, 2001). As the semi-variogram is essential to kriging, a more detail explanation is warranted and given in the next section.

The similarity between two points could be measured by their correlation or covariance; however, both of these techniques require that data values come from the same probability distribution or in other words show stationarity (Clark, 1979). In order to get around this issue, kriging analysis uses the difference between pairs of values rather than single values. These differences in values are classified according to similar vectors and placed into groups (Clark, 1979). In variography, the vectors are called lags and the

groupings are called bins. For instance, all pairs of points separated by 10 m and having a north-south orientation would be grouped into the same bin. Only the average value per bin is plotted in the semi-variogram (Johnston *et al.*, 2001).

A properly chosen lag size is essential as it will aid in revealing any spatial relationships that may be found in the data (Wackernagel, 2003). A small lag size models local variation between neighbouring points more accurately, whereas a big lag size models regional variation more accurately (Johnston *et al.*, 2001).

Before estimations can be produced, the kriging neighbourhood needs to first be defined. The proper definition of the size of the neighbourhood is an important step but not a straight forward one. One way to pick the best model for a given estimation point is to check if adding more points into the calculation lowers the overall variance (Wackernagel, 2003). A much simpler way is to use cross-validation and hold everything constant except the number of neighbours (Wackernagel, 2003).

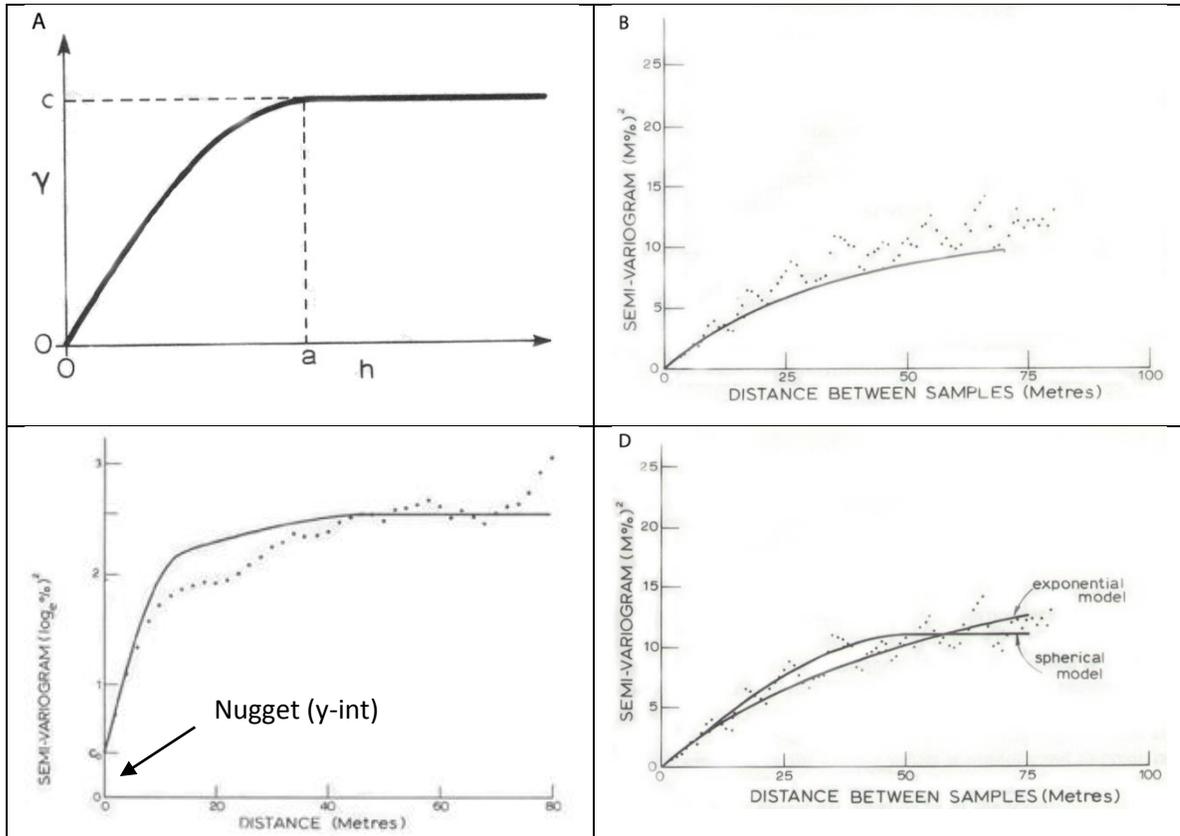
### 2.5.1 The Semi-variogram

The function of the semi-variogram is to model the spatial relationships of a dataset (Johnston *et al.*, 2001). The semi-variogram is half the variogram (hence its name) and it's given by the function

$$\gamma(h) = \frac{\frac{1}{2} \sum [g(x) - g(x+h)]^2}{2} \quad (2)$$

where,  $\gamma$  is the semi-variogram value, which can be understood as a measure of dissimilarity (Johnston *et al.*, 2001);  $h$  is the lag size (i.e., 10 m, N-S orientation);  $g$  is a the grade, concentration or any other value that is to be modeled; and,  $x$  is a location value (i.e. coordinates).

The easiest way to analyze and understand the semi-variogram is by plotting  $h$  on the x-axis and  $\gamma$  on the y-axis (Figure 2.1).



**Figure 2.1.** The semi-variogram and its elements. The lag size ( $h$ ) is plotted on the x-axis and the semi-variogram value ( $\gamma$ ) on the y-axis. The spherical semi-variogram model is shown in A. An experimental (scattered points) and theoretical semi-variogram (best-fit line) are shown in B. C shows the nugget of a semi-variogram. Finally, a comparison between the spherical and exponential models is shown in D. The units of  $\gamma$  are the original units squared. *Source: modified from Clark (1979).*

The spherical or Matheron model (Figure 2.1A) is the ideal shape of the semi-variogram and its role in geostatistics is as essential as the role of the normal distribution in classical statistics (Clark, 1979). Likewise, the experimental semi-variogram and the theoretical semivariogram (Figure 2.1B) are related to each other in the way that a

histogram relates to a probability distribution (Clark, 1979). Furthermore, just as the standard deviation is proportional to the mean in a normal distribution, the sill (C on the y-axis on Figure 2.1A) of the semi-variogram is proportional to the square of the mean of the sample values (Clark, 1979).

The spherical model was derived from theory but it has been shown to have a wide applicability when modeling natural phenomena (Clark, 1979). The spherical model is the best fit for all directions (Johnston *et al.*, 2001). As seen in Figure 2.1A this model rises until it reaches a constant value C, known as the sill. The sill represents a value past which the semi-variance becomes constant, meaning that spatial autocorrelation ceases to be an influencing factor. Also, it has been shown mathematically that the sill is equal to the sample variance (Clark, 1979; Wackernagel, 2003).

The distance at which the sill is reached is known as the range of influence or the range for short (value  $a$  in Figure 2.1A). Figure 2.1C shows another element of the semi-variogram, the nugget or y-intercept. The nugget represents measurement errors, random errors or possible micro-scale variations that are too small to be detected by the semi-variogram (Johnston *et al.*, 2001).

It is important to note that a purely random phenomenon can also have a semi-variogram model. The model in this particular case would be a spherical semi-variogram with a very small range of influence and with a constant sill when  $h$  is greater than zero. In the context of this paper, this is an important notion because the nuggets of Hg and Pb can be compared to determine which pollutant has a more random geographical distribution.

Another important aspect of the semi-variogram is that by definition the axes of  $h$  and  $\gamma$  start at zero, because at zero distance between points there is only one point and the  $\gamma$  of a single point is zero (Clark, 1979).

As seen in Figure 2.1D, the exponential model rises more slowly than the spherical one and it never quite reaches its sill (Clark, 1979). The spherical model is more commonly used than the exponential one in water related studies. For instance, Ouyang *et al.* (2002) used a spherical model to predict the distribution of Pb and Ouyang *et al.* (2003a) used a spherical model to predict the distribution of Hg. Similarly, Little *et al.* (1997) used a spherical model to estimate the levels of contaminant and water variables and so did Gardner *et al.* (2003) to predict stream temperatures. Shi *et al.* (2007) found that spherical models were the best for Hg, Pb, copper and arsenic present in soil samples but that an exponential model was the best for chromium.

Also, it is worth mentioning that the construction of the semi-variogram and the associated predicted values do not depend on the particular statistical distribution of sample points (Clark, 1979). Rather, solely the errors of the predicted values need to be normally distributed in order to calculate unbiased probabilities and produce accurate error maps (Johnston *et al.*, 2001).

Kriging creates a linear estimator to predict values at unsampled locations by using weighted averages of samples values (Clark, 1979). The values that are closer to the location of predicted areas (i.e., cell/block) are given more weight than those located farther away. The semi-variogram is necessary as it allows for the calculation of the variance of a sample and more accurate determination of the contribution of each sample to the final predicted value. In summary, the semi-variogram accounts for the variability

seen in samples, which allows for inferences to be drawn about the phenomena being analyzed.

### **2.5.2 Anisotropy**

The spatial variability of values can be isotropic, meaning that it is the same in all directions or anisotropic, meaning that values have different variability in different directions (Johnston *et al.*, 2001). Anisotropy can be caused by a variety of natural processes such as wind, river flow, runoff and geological structure (Johnston *et al.*, 2001).

Anisotropic semi-variogram models assume that the nugget and sill are the same in all directions but the range varies in an elliptical manner according to a specific direction (Houlding, 2000). This means that the range will be greater for values oriented along the major-axis of the ellipse whereas the range will be shorter for values oriented along the minor axis, effectively creating a semi-variogram envelop (Houlding, 2000).

Another implication of anisotropy is that both distance and direction need to be determined for each estimated value with respect to each sampled value, whereas isotropic models need to determine only the distance (Houlding, 2000).

When dealing with anisotropy, Clark (1979) suggest as a rule of thumb to try search angles at 5, 10, 20 and 45° and that  $h$  should be about 10% of the average distance between samples (ADBS).

## **Chapter 3: Data and Methods**

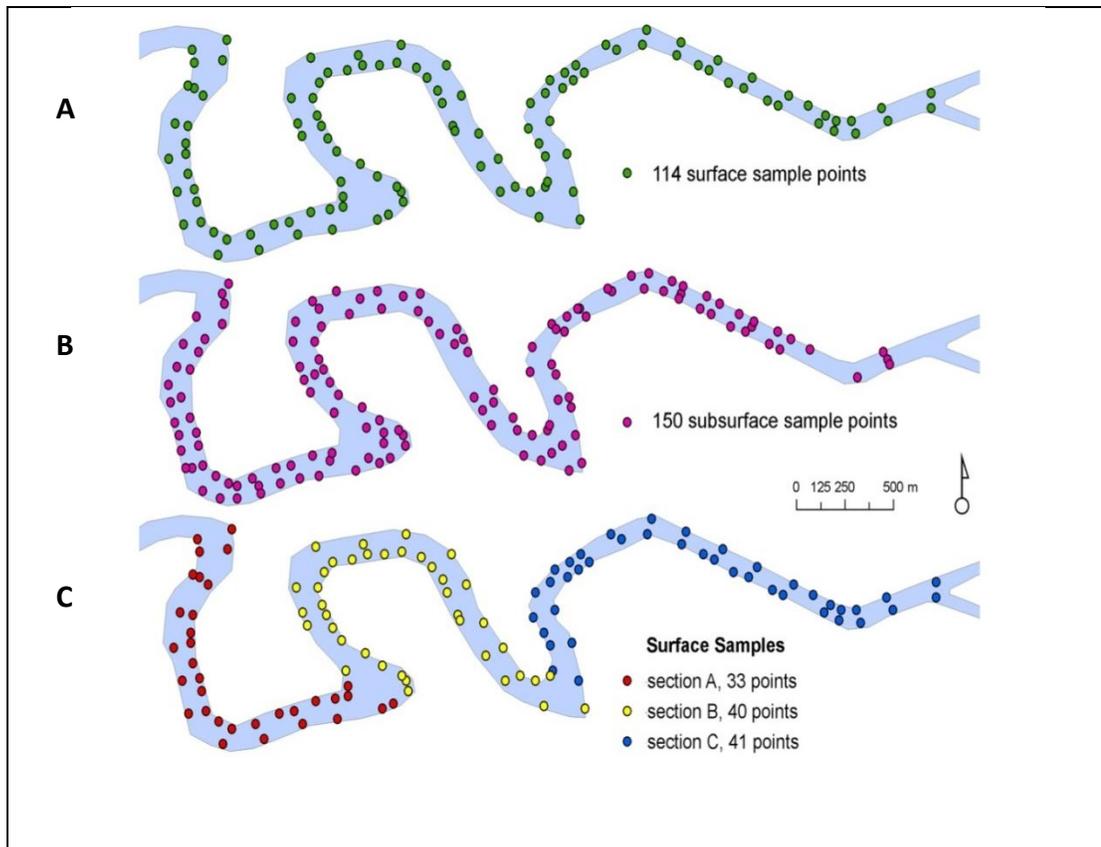
### **3.1 Data**

The data used in this study were collected by the New York State Department of Environmental Conservation in 2005. The original data consist of more than one hundred different pollutants; however, this study only used the data available for Hg and Pb.

One of the goals of NYSDEC's sample design was to collect data in a representative and comprehensive way that would allow for the determination of the nature and extent of sediment contamination within the river. For this reason, the sediment cores were drilled systematically along regularly spaced transects (Sutton, 2006). Some of the samples used in this paper belong to the dredged channel section of the river, which opens a window to assess relatively older and hence unpolluted sediments. In contrast, other samples were taken from the non-dredged stream bank sections of the river, which allows for the analysis of more recently contaminated sediments.

Broadly speaking, the data used in this study were divided in two types, surface and subsurface sediments. Surface sediments mostly consisted of the top six inches (~15 cm) of the core; however, a few 12 inch (~30 cm) composite cores were also part of the dataset. Subsurface sediments consisted of all other samples deeper than one foot (>30 cm). Surface samples emphasize the current (2005) and ambient conditions of the biological active strata, whereas subsurface samples emphasize the history of pollutant depositional patterns (Sutton, 2006). This study used 114 surface and 150 subsurface samples from a total of 182 cores, meaning that a single core can have one or more data

values associated with it. The geographical distribution of the surface sample points can be seen in Figure 3.1A and that of subsurface ones can be seen in Figure 3.1B.



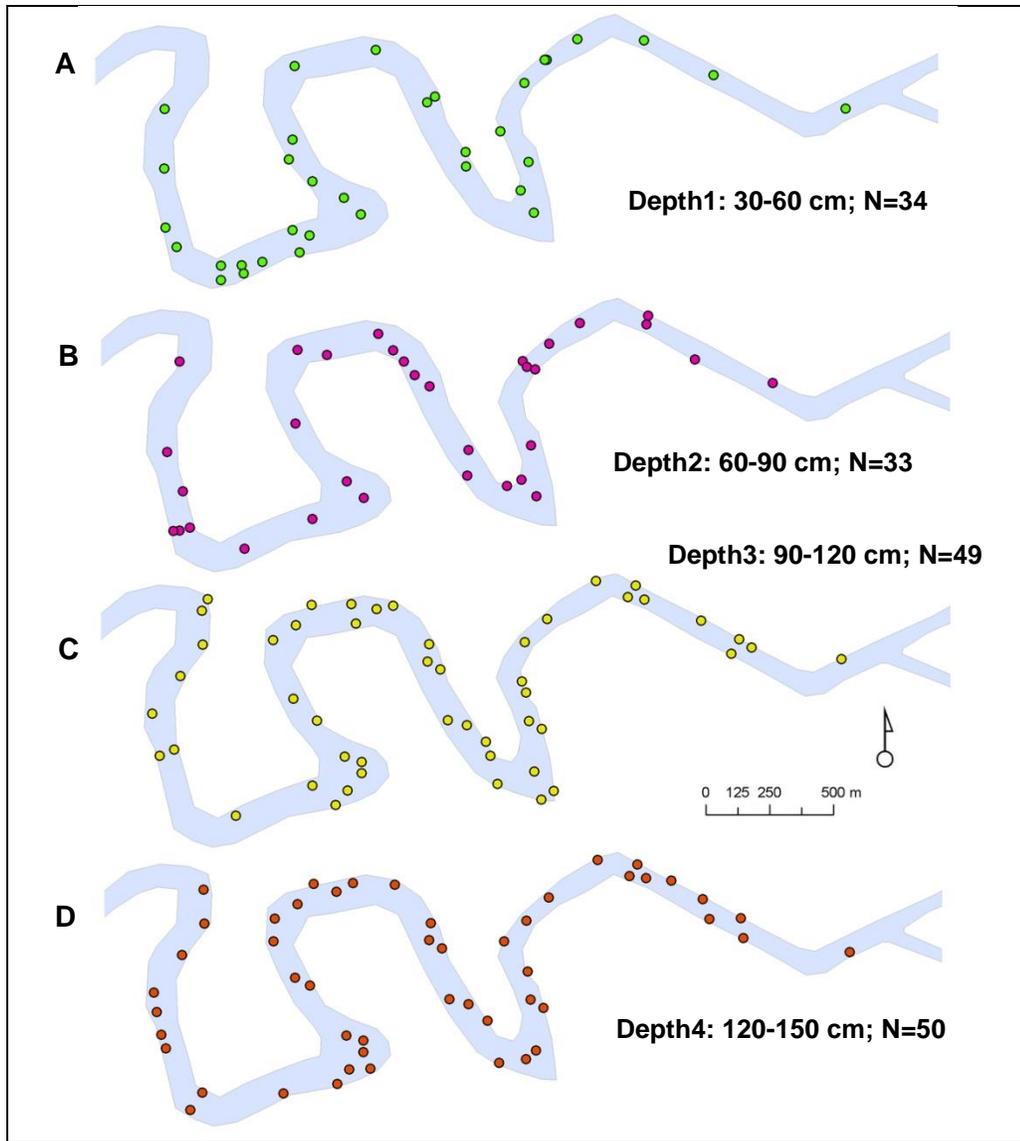
**Figure 3.1.** Distribution of sample points within the study area. There are two broad types of sediment data, surface (A) and subsurface sediments (B). As seen in C the surface sediments were further divided into sections A, B and C.

The sample size used in this paper was smaller than the original amount of sample points gathered by NYSDEC. This reduction occurred because null value samples (i.e., sediment loss) were taken out of the analysis. Similarly, sample points that were located too far away from the section of the river studied in this paper were removed.

Preliminary analysis showed that the kriging search window (i.e., kriging neighbourhood) included sample points that were spatially close but separated by dry land. In order to test whether this issue affected the accuracy of the prediction maps, the original surface sediment data set of 114 points (Figure 3.1A) was divided into subsets

(Figure 3.1C). The criteria used to divide the original data set were the need for a robust dataset of at least 30 observations and the splitting up of close meanders (in order to assign the correct weight to neighbouring sample points). As a result, three study area sections were created: section A with 33 samples, section B with 40 samples, and section C with 41 samples.

Figure 3.2 shows the distribution of subsurface sample points at four depth levels. Each depth level has a different sample size as not all sediment cores were sampled at the same depth intervals. Depth 1 (30-60 cm) has 33 samples, Depth 2 (60-90 cm) has 34 samples, Depth 3 (120-150 cm) has 49 sample points, and Depth 4 (150-180 cm) has 50 samples. The depth interval in centimetres is an approximation of the original units, which were recorded in feet by NYSDEC. In terms of subsurface samples, only these four datasets were used to carry out ordinary kriging analyses.



**Figure 3.2.** Distribution of data points within four depth levels in the study area.

It should be noted that the data values used in this study were obtained by taking a small amount of sediment from each core which was analyzed in a laboratory. The pollutant concentration values found were then assumed to be representative of the whole core interval. This assumption is safe as the sediment core interval was relatively homogeneous in its amount of sand, silt and clay, as different types of particles can hold different amounts of pollutants (Smol, 2008). More details of the sampling design can be found in the Buffalo Sediment Study authored by Sutton (2006).

The original spatial data gathered by NYSDEC were in the form of shapefiles that used the NAD 1983 datum and the State Plane New York West FIPS 3103 (units: feet) projected coordinate system. In this study, these shapefiles were re-projected to the Universal Transverse Mercator Zone 17N (units: metres) coordinate system to enable direct importation with other on-going research projects.

### **3.2 Exploratory Data Analysis**

Exploratory data analysis (EDA) was carried out using surface and subsurface data at depths ranging from 30 to 300 cm. As not all sediment cores were sampled at the same depth intervals, the sample size varied according to depth.

The EDA that was carried out resulted in the creation of histograms and bar graphs for surface and subsurface sediments. These are standard procedures that have been used in similar studies to explore the nature of Hg and Pb pollution (Ouyang *et al.*, 2003a; Fabbri and Trevisani, 2005). Moreover, the data were further explored by carrying out correlation scatter plots between Hg and Pb. EDA was done using R 2.8.1 statistical software and MS Excel 2007.

### **3.3 Log-transformation**

There is no agreement on whether non-normal distributions should be transformed prior to fitting the semi-variogram and performing kriging. Some authors argue that kriging, requires a normal distribution of values and thus, non-normal distributions need to be transformed (Houlding, 2000; Johnston *et al.*, 2001). Forsythe *et al.* (2004) suggest that statistically insignificant models (average standard error > 20) may become significant if

positively skewed data are log-normalized. When ASE values are  $>20$ , predictions stray quite far from the original data point values. Similarly, Ouyang *et al.* (2003a) argue that even though normality is not essential for kriging, the presence of pronounced skewness in a data set warrants some kind of transformation. Likewise, Clark (1979) argues that kriging as an interpolator does not require normality rather only kriging prediction errors rely on normality. Conversely, Wackernagel (2003) warns against the use of lognormal transformation as this procedure causes the mean to shift giving poor results. There is more agreement, however, with respect to the use of lognormal as the best choice for transforming non-normal data (Clark, 1979; Houlding, 2000; Wackernagel, 2003) as long as there are no zero values in the dataset

In this paper, log-transformations of non-normal distributions were done using base ten log. Only logged values were used for the fitting of theoretical semi-variograms. Log calculations and transformations were performed in MS Excel 2007 and these logged values were then used in the ArcGIS environment.

### **3.4 Ordinary Kriging Analysis**

This study carried out ordinary kriging analysis using ArcGIS Geostatistical Analyst 9.3. By default this software fits a spherical semi-variogram model and calculates optimal values for the lag size ( $h$ ), number of lags, nugget, partial sill (difference between sill and nugget) and the range's major and minor axes (Johnston *et al.*, 2001). Furthermore, in terms of the kriging neighbourhood, the default options set the number of neighbours to five with the inclusion of at least one neighbour and a search ellipse with axes crossing at  $45^\circ$  angles. These kriging neighbourhood default parameters were used throughout all

kriging analyses carried out in this study. A combination of only default options was used as the first kriging model for each data set.

In order to carry out ordinary kriging analysis, the present study used ArcGIS default parameters, literature derived parameters and experimental parameters derived from preliminary analysis. Therefore, non-default theoretical semi-variogram models and parameter values were tested. For instance, anisotropy was tested with search windows at 5, 10, 20 and 45° and h equal to 10% of ADBS after Clark (1979). Similarly, the lag size and lag number were modified to incorporate Wackernagel's (2003) suggestion to only use (in the semi-variogram) pairs of points up to a distance of half the diameter of the region being analyzed. These parameters in combination with default ones formed the basis for other kriging treatments.

ArcGIS Geostatistical Analyst has eleven different theoretical semi-variogram models available to choose from. This paper only tested spherical, exponential and Gaussian models. Also, this study tested the effect of removing trends from the data sets. Only none, first and second order trend removals were tested. When a trend is present in a dataset, this can be removed right inside the Geostatistical Analyst. In this case the semi-variogram is fitted using the residuals but the trend is placed back into the calculation right before calculating the final values (Johnston *et al.*, 2001).

The present study conducted ordinary kriging analysis using the surface sediment data set (Figure 3.1A), three surface sediment data sets corresponding to Sections A, B and C (Figure 3.1C), and subsurface sediments data sets (Figure 3.2). Ordinary kriging of subsurface sediments was performed at four different depth levels in an attempt to analyze a 3D phenomenon (i.e., sediment contamination) as several 2D layers. This

approach was derived from Ouyang *et al.* (2003a), who kriged the Cedar-Ortega water system at 0.1, 0.5, 1.5 and 2.0 metres. Similarly, Fabbri and Trevisani (2005) analyzed soil Pb pollution levels at eight different depths.

Sixteen different kriging models per data set were carried out. The most accurate kriging model for each data set was chosen using the Best Model Index (BMI). More details on this index are given in the next section. Prediction maps and prediction error maps for the best Hg and Pb kriging models were created. A full list of the best kriging models appears in Appendix A.

The prediction maps for sections A, B and C were spliced together in order to produce a map that covered the original data extent. The maps created by joining Sections A, B and C are referred to as spliced maps. In contrast, the maps created by using the complete data set (n=114) of the study area are referred to as standard maps.

This paper also used ordinary kriging to produce a prediction surface of water depth along the study area. The reason for this was to facilitate the discussion concerning the geographical distribution of Hg and Pb. In this particular case, the default values of ArcGIS Geostatistical Analyst were used.

The symbology of the prediction surfaces was adjusted to create TEL and PEL isolines for Hg and Pb. These isolines were produced using the values shown in Table 2.1 and added to the prediction maps.

### **3.5 Model Cross-Validation**

ArcGIS Geostatistical Analyst allows for kriging analysis in combination with cross-validation. The latter process results in the generation of five prediction error parameters:

mean (M), mean standardized (MS), root mean square (RMS), average standard error (ASE) and the root mean square standardized (RMSS).

Cross-validation is a simple way to compare the type of semi-variogram, its associated parameters and the kriging neighbourhood among different kriging models. This procedure entails the removal of one sample value from the dataset and the estimation of a value at this very point with n-1 samples (Wackernagel, 2003). This means that different models can be compared on the basis of how well they estimate measured sample values (Wackernagel, 2003). The difference between sampled values (SV) and estimated values (EV), gives an indication of how well the predicted value fits in the neighbourhood of sampled values (Johnston *et al.*, 2001).

Cross-validation computes the M of all these differences (i.e. SV- EV). If M is close to zero then there is no bias, if M is a negative then the kriging model overestimates the real value and if M is positive then it underestimates the real value (Wackernagel, 2003). The MS is computed by dividing M by the standard deviation and it is more useful than M when comparing models as M is dependent on scale (Johnston *et al.*, 2001).

Cross-validation also computes the RMS, which is found by first squaring the differences between SV-EV, then calculating the average of these differences and finally, taking the square root of this average. This means that the RMS, contrary to M, can never be negative.

The ASE is calculated by dividing each predicted error by the standard deviation obtained with n-1 samples (Wackernagel, 2003). If the ASE is close to the RMS, then the model is assessing the variability in the prediction appropriately. If ASE is greater

than RMS then the model overestimates the true variability and if ASE is lower than RMS then the model underestimates the variability (Johnston *et al.*, 2001). The RMS and ASE are important as they allow for the comparison of different models that use the same data (Johnston *et al.*, 2001; Wackernagel, 2003).

The last prediction error measurement is the RMSS, which is equal to the square root of the sum of the squared predicted errors divided by the squared standard deviation and divided by the total number of samples (Wackernagel, 2003). If the RMSS is close to one, then the model and its parameters are adequate as the n-1 estimation error is equal to that predicted by the model (Wackernagel, 2003). As it is in the case of M, if the RMSS is greater or lower than one, under or over estimation of the variability of the predicted values occurs (Johnston *et al.*, 2001).

In general terms the best kriging model is that which has the MS closest to zero, the smallest RMS, the ASE closest to the RMS, and the RMSS closest to one (ESRI, 2001). These four prediction error parameters served as the basis to build the best model index (BMI) and hence, choose the best kriging models to produce prediction maps.

Another way to check for the quality of a model is to look at two graphs produced by the Geostatistical Analyst. The predicted values graph, plots measured values against predicted values. If a model is good then the predicted values should scattered around the 1:1 line (Johnston *et al.*, 2001). The slope of the line of best fit is usually less than one as kriging tends to under predict large values and over predict smaller ones (Johnston *et al.*, 2001). The second useful graph is the QQ plot. As the error of a robust model should be independent, then the residuals shown in a QQ plot should be close to the 1:1 line

(Johnston *et al.*, 2001). In this case mapping methods that rely on normality such as the prediction error map can be used confidently.

### 3.6 The Best Model Index (BMI)

The relatively large number of kriging models produced in this study called for an objective way to compare them and choose the best one for mapping. As a result, a best model index (BMI) was created.

The BMI uses the MS, RMS, ASE and RMSS obtained through cross-validation and arranges them in a way that meets the “good kriging model” criteria given in the literature. The BMI is given by the following formula,

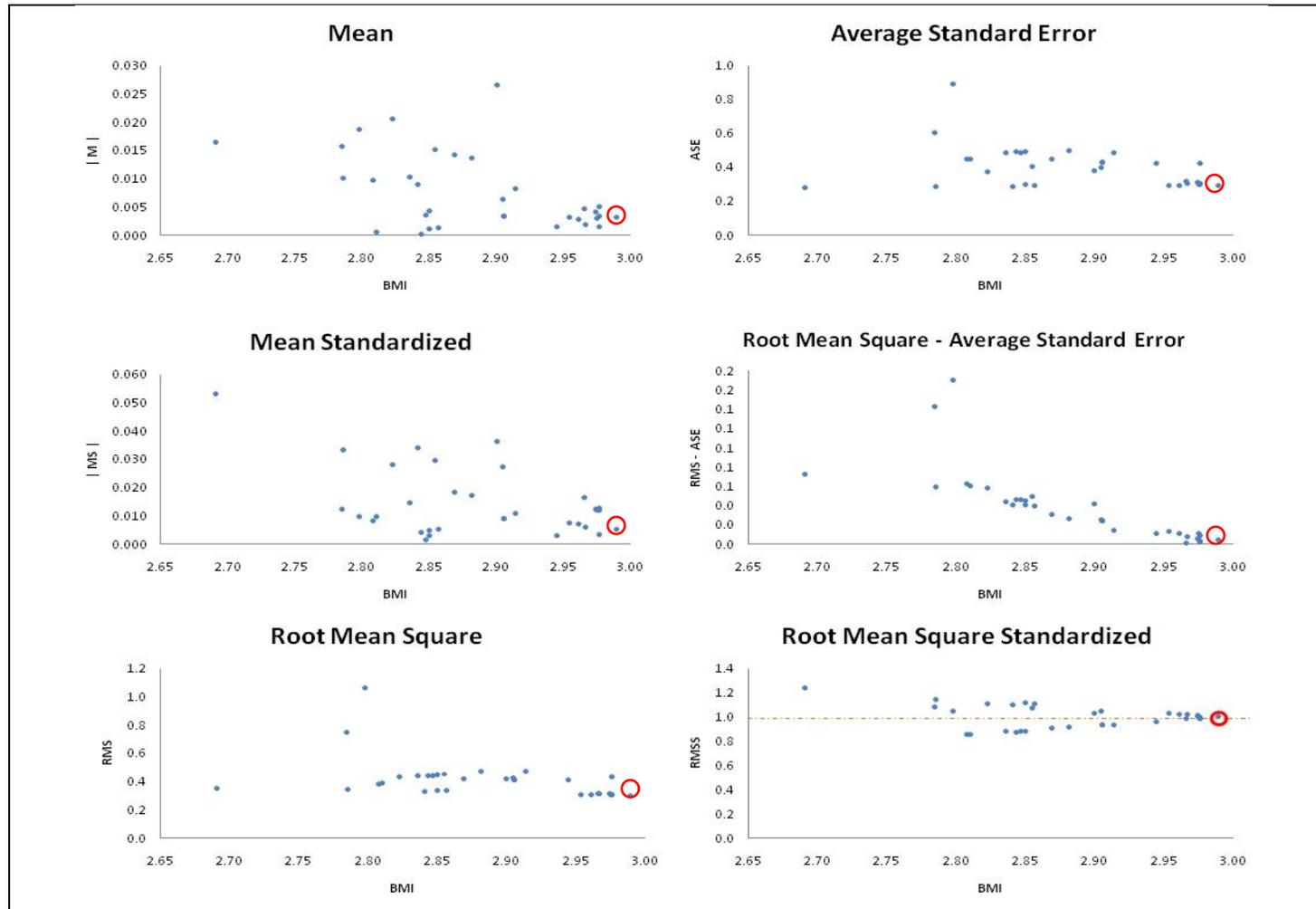
$$BMI = \frac{1}{1 + |MS|} + \frac{1}{1 + (|RMS - ASE| * (RMS * ASE))} + \frac{1}{1 + |1 - RMSS|} \quad (3)$$

In theory, a model could score a maximum BMI value of three, which is the best score possible. This will occur when the MS is zero, the difference between its RMS and ASE is zero, and its RMSS is equal to one. This perfect model is very unlikely if not impossible to produce as can be seen in Figure 3.3, which shows the relationship between the predicted errors measurements and the BMI.

The scatter plots shown in Figure 3.3 were created using prediction errors obtained from preliminary analyses using 30 different Hg kriging models. Figure 3.3 shows that models with the highest BMI also tend to have a low M, MS, RMS, ASE, MSE-ASE, and RMSS close to one (red circles in Figure 3.3). Also, Figure 3.3 shows that while some models have good scores for a particular error measurement; these don't achieve a good score in all the measurements. Out of all the models tested during preliminary analysis no model scored the best in each of the four error categories.

The model with the highest BMI is not necessarily the model with the lowest M, RMS, ASE and RMSS closest to one; rather it is the model that does best overall among all four error measurement categories. For instance, the kriging model that scored the highest BMI in the preliminary analysis was one that scored the eighth lowest MS, the lowest RMS, the smallest difference between RMS and ASE, and had an RMSS closest to one. Consequently, this best overall score (i.e. highest BMI) was assumed to be indicative of the best kriging model.

It should be noted that the BMI was derived from prediction error measurements of log-transformed values. It is possible that this index may not perform the same when using non-transformed data. Furthermore, the BMI can only be used to compare models that use the same data set (i.e., same variable and sample size). Finally, it is worth mentioning that a more robust index than the BMI is present in the literature (Barabas *et al.*, 2001); however, it is more complex and goes beyond the scope of the present study.



**Figure 3.3.** Scatter plots of prediction error measurements against the best model index (BMI). The model with the highest BMI consistently scored among the lowest values on all error categories tested. The kriging model that scored the highest BMI was used to create the prediction maps for Hg and Pb concentrations in the study area.

### 3.7 Ordinary Kriging Flow Chart

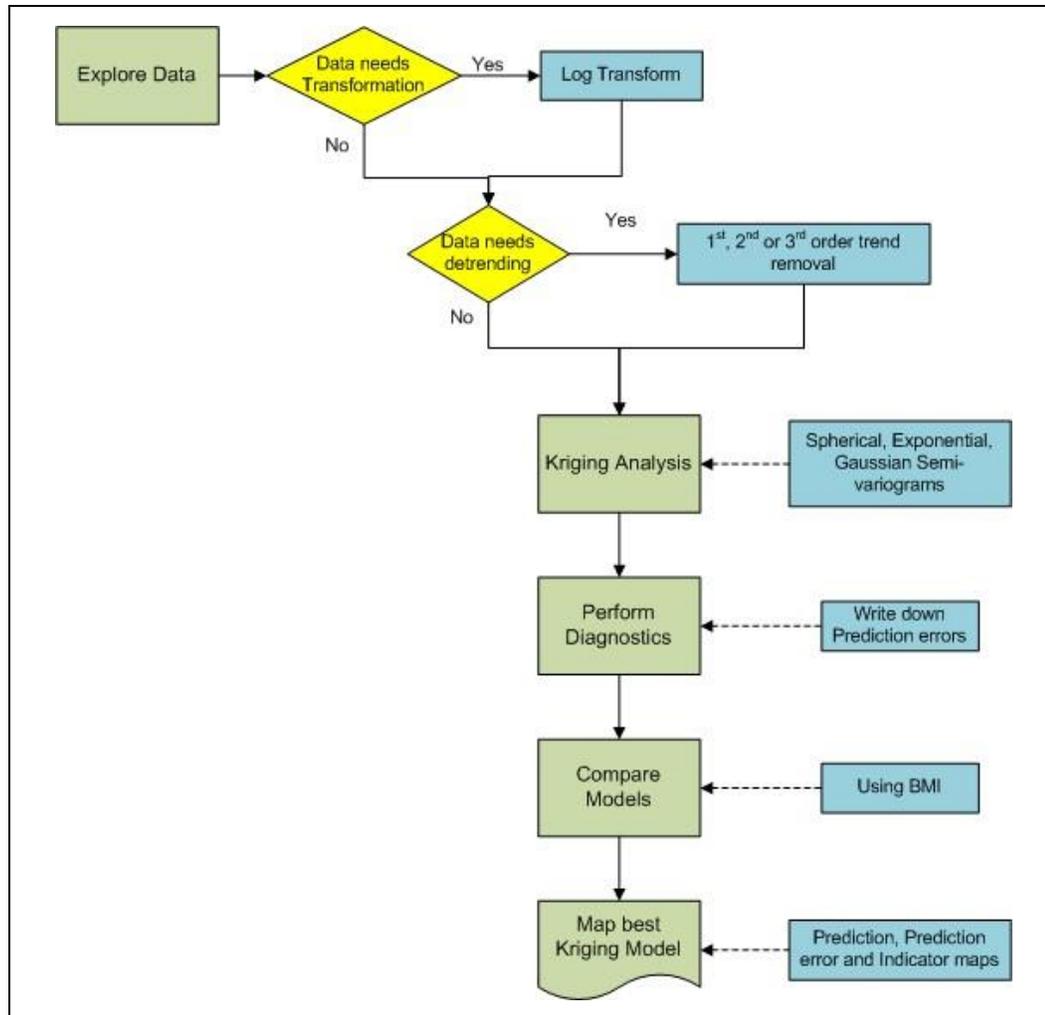
This paper followed the general work flow proposed by Johnston *et al.* (2001) and which can be seen in a modified format in Figure 3.4. The very first step was to explore the data to see if the values were normally distributed and to see if there were any trends. If the data set was non-normal, it was log-transformed using log base 10. (Note: preliminary analysis showed that the data were positively skewed and that all values were bigger than 0; hence, the use of log-transformation was justified.) Several kriging models were then tested and their prediction errors were recorded. These errors were then used in the BMI to determine the best model out of at least 16 different kriging models. Finally, the best models were used to create prediction maps and prediction error maps.

A systematic approach was used to create at least 16 kriging models per data set. For a given contaminant or data set the first kriging model was obtained by using the Geostatistical Analyst's default parameters. The second and third models used the same default parameters except that the default spherical semi-variogram model was changed to exponential and Gaussian, respectively.

The fourth and fifth kriging models included 1<sup>st</sup> and 2<sup>nd</sup> order trend removals using the best of models 1 to 3. The sixth model tested the presence of anisotropy using the best from models 1 to 5. The seventh to tenth kriging models tested the effect of changing the default lag size and number of lags. These parameters were changed according to values that were obtained directly or indirectly from the literature. For instance, number of lags equal 10 was taken directly from Ouyang *et al.* (2003) whereas lag size of 76 m was deducted from the same study. Further combinations of default and

non-default parameters were tested until at least 16 different kriging models were produced.

The standard prediction maps for Hg and Pb were created using the 114 surface sediment samples available and were classified using the quantiles classification method. Quantiles classification allows for comparison between different pollutants because each class will always have values. The prediction and prediction error map values were divided into ten and five classes respectively. The former number allowed for the visualization of subtle changes, whereas the latter number allowed for quick prediction error comparisons between pollutants. These surface classification schemes were used as the reference to manually classify all other maps so that comparison between different maps could be done sensibly. The range of values of the prediction standard error maps were obtained by using the internal log transformation function available in ArcGIS Geostatistical Analyst whereas that for the prediction maps were obtained by simply using the exponential function (i.e.,  $10^X$ ).



**Figure 3.4.** Ordinary kriging methodology flow chart. *Source: modified after Johnston et al. (2001)*

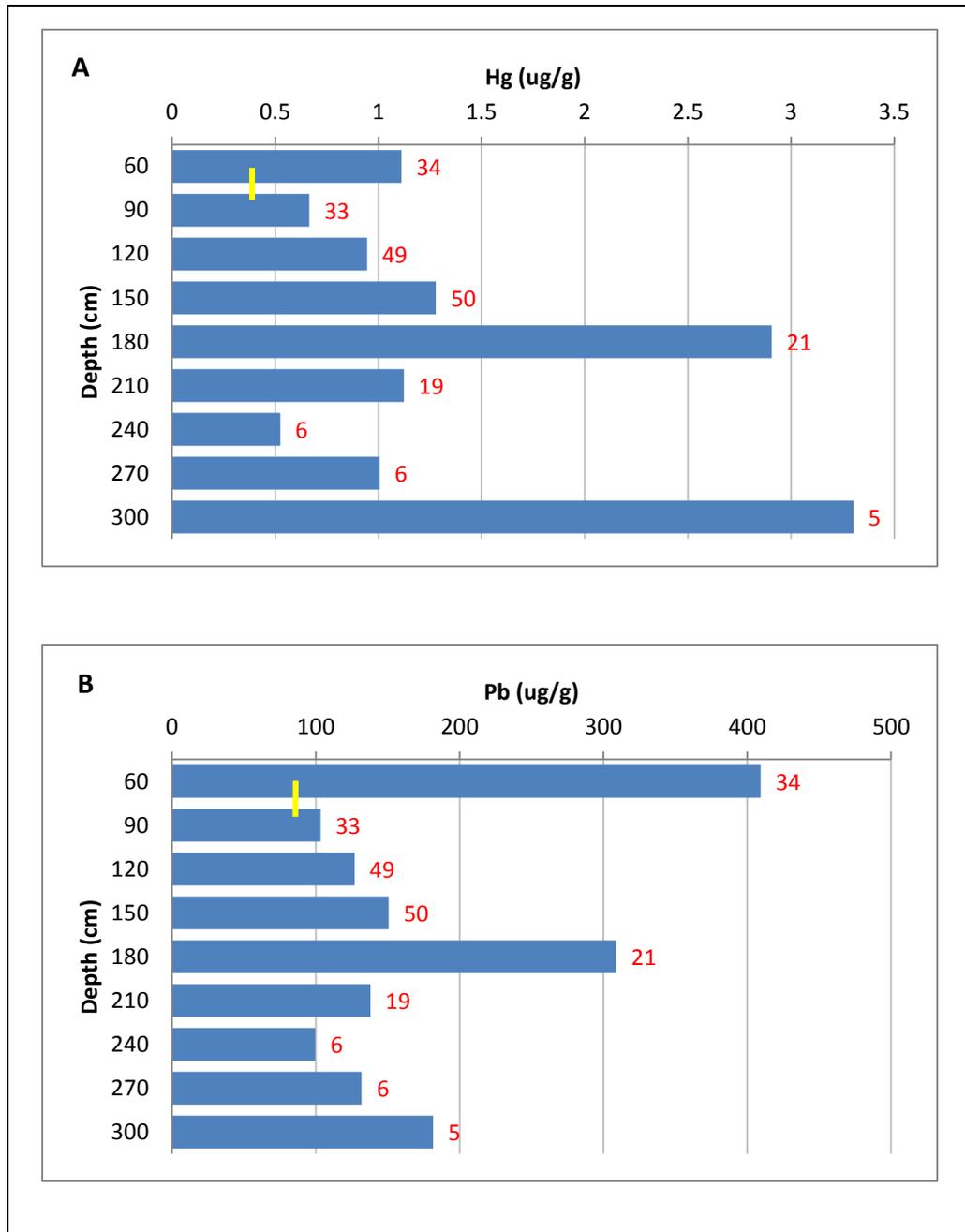
## Chapter 4: Results

### 4.1 Exploratory Data Analysis

Hg and Pb needed to be logged because their histograms showed a positively skewed distribution. This meant that both these metals had a relatively large proportion of low values and a smaller proportion of higher values.

The skewness index values of both metals lowered considerably after log transformations were applied. For instance, the index value for Hg lowered from 6.28 to 1.15, whereas that for Pb lowered from 9.25 to 1.77. However, some skewness was left in the data as the index values are above one. In short, log transformation reduced the skewness in the data for both metals but did not make their distribution entirely normal.

The concentration of Hg and Pb in subsurface samples differs with respect to each pollutant and with respect to depth. Figure 4.1 shows that the highest average concentration of Hg was found in the bottom most layer (4.1A), whereas the highest average concentration of Pb was found in the top most layer (4.1B). Overall, there seems to be a low to high trend as a function of depth for Hg, whereas an opposite trend seems to be the case for Pb. Moreover, at the 180 cm depth class there was a high concentration for both Hg and Pb, which is strikingly higher than the immediate lower and higher depth classes. Furthermore, an interesting pattern is that the first four depth classes for Hg and Pb are proportional to each other. The yellow line on the 60 cm depth class bars of Figure 4.1 represents the average value of surface sediments (0 – 30 cm) for Hg and Pb.



**Figure 4.1.** Average concentration of Hg and Pb in subsurface sediments. The mean concentration value for each pollutant is plotted with respect to depth (30 cm intervals starting at 30 to 60 cm). The yellow line on the top bar shows the mean of surface sediments as a reference (Hg = 0.288  $\mu\text{g/g}$ ; Pb=86.16  $\mu\text{g/g}$ ). The number at the right end of each bar indicates sample size.

Table 4.1 shows the summary statistics for eight different data sets: four surface and four subsurface ones. All surface sediment data sets represent those found between 0-30 cm, whereas the subsurface sediment datasets represent the first four depth classes shown in Figure 4.4.

Table 4.1 reveals that Hg for Section B had the highest mean (0.565  $\mu\text{g/g}$ ) and variability (standard deviation=1.336  $\mu\text{g/g}$ ) of all surface samples, even higher than that for the entire study area. In contrast, Section A has the highest median value (0.095  $\mu\text{g/g}$ ) and the lowest maximum value (0.48  $\mu\text{g/g}$ ).

As seen in Table 4.1, Hg subsurface sediments are much more contaminated than surface ones. For instance, the four highest Hg mean values are all found in subsurface samples. In fact, the lowest Hg mean in subsurface samples is about three times higher than the highest mean of surface samples.

Depth 4 shows the greatest skewness of all subsurface datasets, as not only does it have the highest Hg mean value, maximum value and standard deviation but also the second lowest median value. Furthermore, it is worth noting that Depth 2 has the highest median value (0.3  $\mu\text{g/g}$ ) but the lowest maximum value (6.9  $\mu\text{g/g}$ ) of all subsurface data sets.

In terms of Pb in surface sediments, Table 4.1 shows that section C has the highest mean (113.80  $\mu\text{g/g}$ ) and standard deviation (405.51  $\mu\text{g/g}$ ) of all study area sections, higher than even the entire study area. Section A has the highest Pb median value and also the lowest maximum value. Similar to Hg, subsurface sediments show a higher Pb concentration than surface sediments. However, in this case the magnitude of

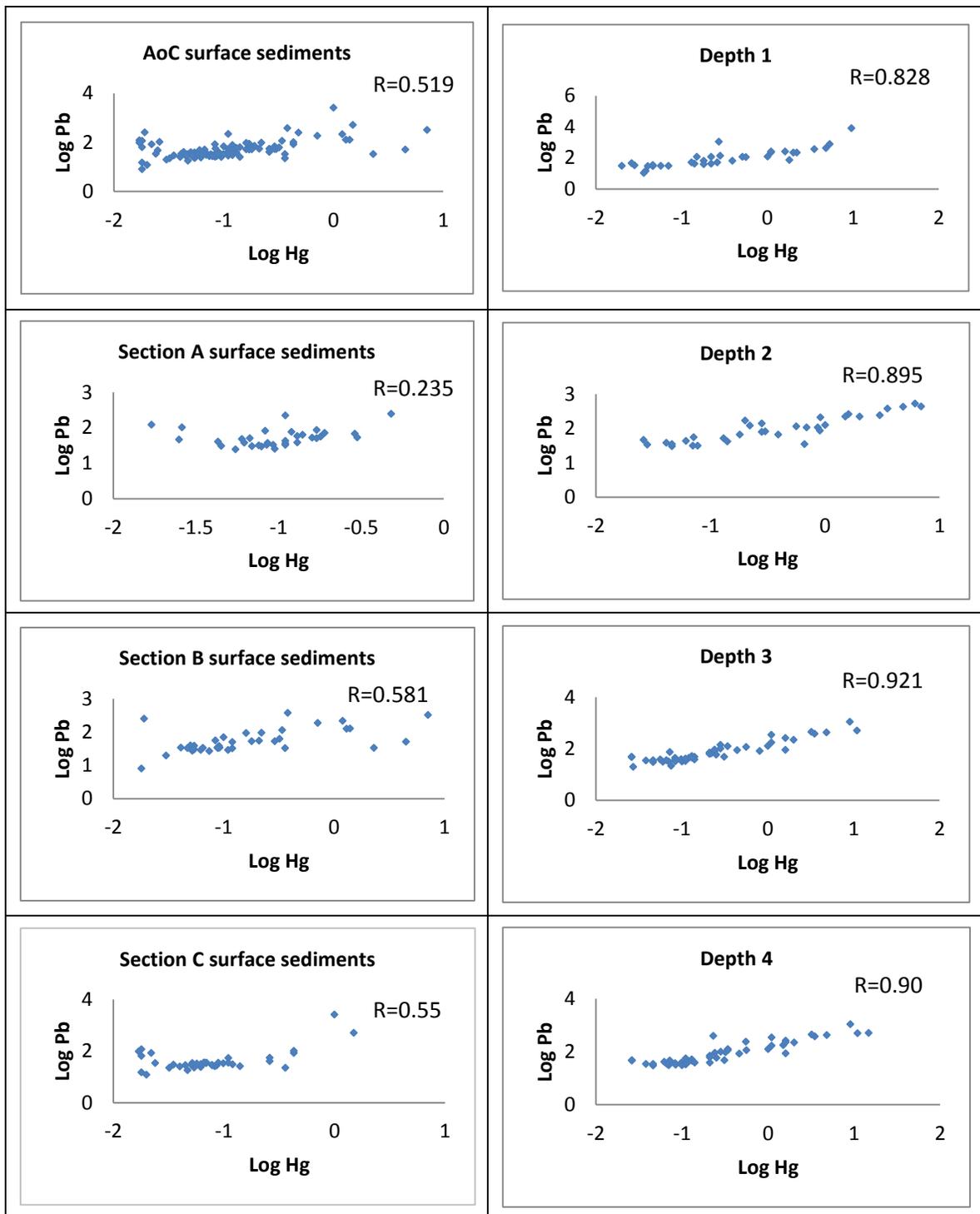
the difference between Pb in surface and subsurface sediments is much less dramatic than that of Hg.

Just as with Hg, Depth 2 has the highest Pb median value (87.4  $\mu\text{g/g}$ ) and the lowest maximum value (538  $\mu\text{g/g}$ ). In contrast to Hg, the Depth 1 has the highest Pb mean value (409.3  $\mu\text{g/g}$ ) and the highest standard deviation (1,449.7  $\mu\text{g/g}$ ) of all subsurface sediments.

**Table 4.1.** Summary statistics of surface and subsurface sediment data sets. Ordinary kriging analysis was carried out on each of the eight data sets shown.

Dataset	n	Hg					Pb				
		Mean	Std. Dev.	Min.	Median	Max.	Mean	Std. Dev.	Min.	Median	Max.
Surface samples											
Study area	114	0.288	0.829	0.017	0.086	7.1	86.16	248.67	8.1	35.8	2600
Section A	33	0.120	0.092	0.017	0.095	0.48	63.35	50.66	25.0	48.8	252
Section B	40	0.565	1.336	0.018	0.092	7.1	76.65	84.06	8.1	38.1	380
Section C	41	0.153	0.277	0.017	0.066	1.5	113.80	405.51	12.1	32.5	2600
Subsurface samples											
Depth1	34	1.112	1.991	0.020	0.240	9.5	409.30	1449.65	10.9	71.1	8510
Depth2	33	1.138	1.771	0.026	0.300	6.9	142.80	134.71	30.4	87.4	538
Depth3	49	0.946	2.122	0.026	0.210	10.9	127.00	185.36	19.8	58.7	1100
Depth4	50	1.279	2.853	0.026	0.220	14.7	150.60	194.04	30.4	60.5	1100

Figure 4.2 shows the scatter plots of log Pb against log Hg by data set and sediment type. The correlation in all cases is positive, with subsurface sediments showing stronger correlation than surface ones. All correlations are significant ( $p < 0.01$ ) except for Section A. Section B shows the strongest correlation among surface sediments with an R value of 0.58, whereas Depth 3 shows the strongest correlation among subsurface sediments with an R value of 0.921.



**Figure 4.2.** Correlation scatter plots of log Pb against log Hg. Section B and Depth 3 show the strongest correlations for surface and subsurface sediments respectively. All R values are significant ( $p < 0.01$ ) except for section A ( $p > 0.1$ )

## 4.2 Ordinary Kriging Analysis

The results for kriging analysis were different for Hg and Pb, both in terms of their best kriging models and in their geographical distributions within the study area. The 16 chosen models (i.e., highest BMI) together with their respective parameters are shown in Table A.1 in Appendix A.

Models that used the Geostatistical Analyst default parameters scored the highest BMI only in one instance, that for Hg in Section C. Models that used a Gaussian theoretical semi-variogram with anisotropy scored the highest BMI with the study area dataset for both, Hg and Pb. Similarly, the same semi-variogram model was the best for Pb in Depth 3 and Depth 4. An isotropic Gaussian semi-variogram model was the best for Hg and Pb in Depth 2. Isotropic spherical and exponential semi-variograms produced five of the eight surface sediment kriging models. A final tally shows that Gaussian was used in seven of the best kriging models, followed by exponential in five models and spherical in four models.

Accounting for anisotropy produced the best model in about half of all the kriging models. When an anisotropic model scored the highest BMI, the direction of the search ellipse had a range between 273 to 294 degrees. Only in the case of Pb in Section B was the angle out of this range (321degrees).

A second order trend removal was required for modeling Hg in Section B and Depth 1; whereas a 1<sup>st</sup> order polynomial was the best for Hg at Depths 3 and 4. In terms of Pb, only 1<sup>st</sup> order trend removal was required for Section B, Depth 2 and Depth 3. In contrast, the dataset for the entire study area required no detrending. In total, seven out of sixteen kriging models accounted for a trend in the data.

The semi-variogram for Hg showed nuggets larger than those of Pb in six out of the eight data sets. For the whole study area, the nugget of Hg was almost twice that of Pb. Similarly, the nuggets of Hg were almost double those of Pb among all depth levels. However, when the study area was divided into sections, Hg showed a higher nugget than Pb only in Section C.

The sill was higher for Hg than for Pb in all surface sediments. The highest sill ( $0.409 \mu\text{g}/\text{g}^2$ ) was found in Section B for Hg and in Section C for Pb ( $0.130 \mu\text{g}/\text{g}^2$ ). These high sill values parallel the findings of Table 4.1, where Hg and Pb showed the highest standard deviations in Section B and Section C, respectively.

The range for Hg and Pb in the study area was 754 m. Not surprisingly, the ranges of all study area sections were smaller, varying from 211 m to 496 m for Hg and from 77 m to 694 m for Pb. This difference in range size between the study area and its sections was expected as the area covered by the data points in each of the sections is smaller than that of the study area. For subsurface sediments the range of Hg varied from 498 m to 900 m, whereas for Pb the ranges varied from 529 m to 754 m.

In terms of lag sizes, 76 m was the best size for Hg in three out of the four data sets that had equivalent study area sizes (i.e. study area and Depth 1 to 4 data sets). Similarly, 76 m was the best lag size for Pb in all equivalent study areas. In the study area sections, Hg required a lag size that ranged from 22.6 to 50 m; whereas, Pb required a lag size between 17.5 and 70 m.

The semi-axis for Pb (1,290 m) in the study area was much longer than that for Hg (754 m). The same semi-axis value was also the best for Pb at Depths 3 and 4. The

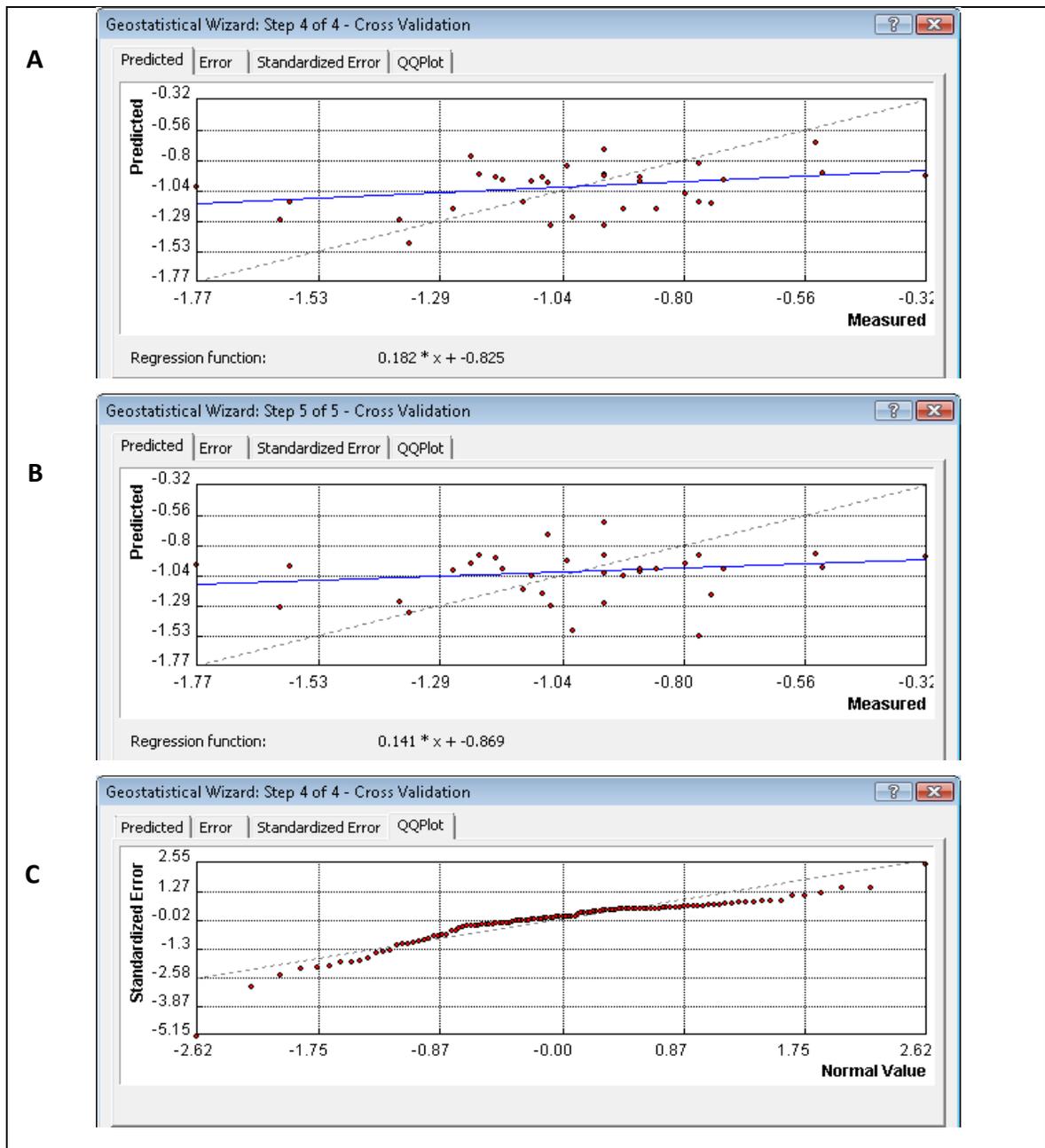
semi-axis for Hg in subsurface sediments varied significantly, ranging from 497.9 m (Depth 1) to 1,290 m (Depth 4)

It is noteworthy that the kriging model that scored the best BMI for Hg was not necessarily the best for Pb. Moreover, there was no single set of parameters that was overwhelmingly the best in all circumstances even within Hg or Pb kriging models.

Also, it is worth mentioning that although some data sets showed a clear trend when analyzed with the trend analysis tool of the Geostatistical Analyst, accounting for this trend did not necessarily produce the best kriging model. For instance, a strong parabolic trend was seen in Section A but a plain model produced better results than a detrended model.

In terms of the prediction error and Q-Q plots, the worst and best kriging models did not show major differences. For instance, using the plots for Hg Section A (Figure 4.6) as an example, it can be seen that the best (Figure 4.3A) and worst kriging models (Figure 4.3B) are very similar. The slope of the best model is only slightly higher than that of the worst model and the error term is only slightly lower. The line of best fit (blue line) is far from lining up with the 1:1 line. In fact, in no single case, was the slope of the best kriging model bigger than 0.5.

A constant characteristic of all kriging treatments was their pronounced deviation from the 1:1 line of the Q-Q plots at low and high values (Figure 4.3C). Only at mid values did the residuals line up along the 1:1 line.



**Figure 4.3.** Prediction error and Q-Q plots. The best model for Hg in Section A is shown in A and the worst model for Section A is shown in B. Also, a Q-Q plot for hg in Section A is shown in C.

### 4.3 Distribution of Hg and Pb in surface sediments

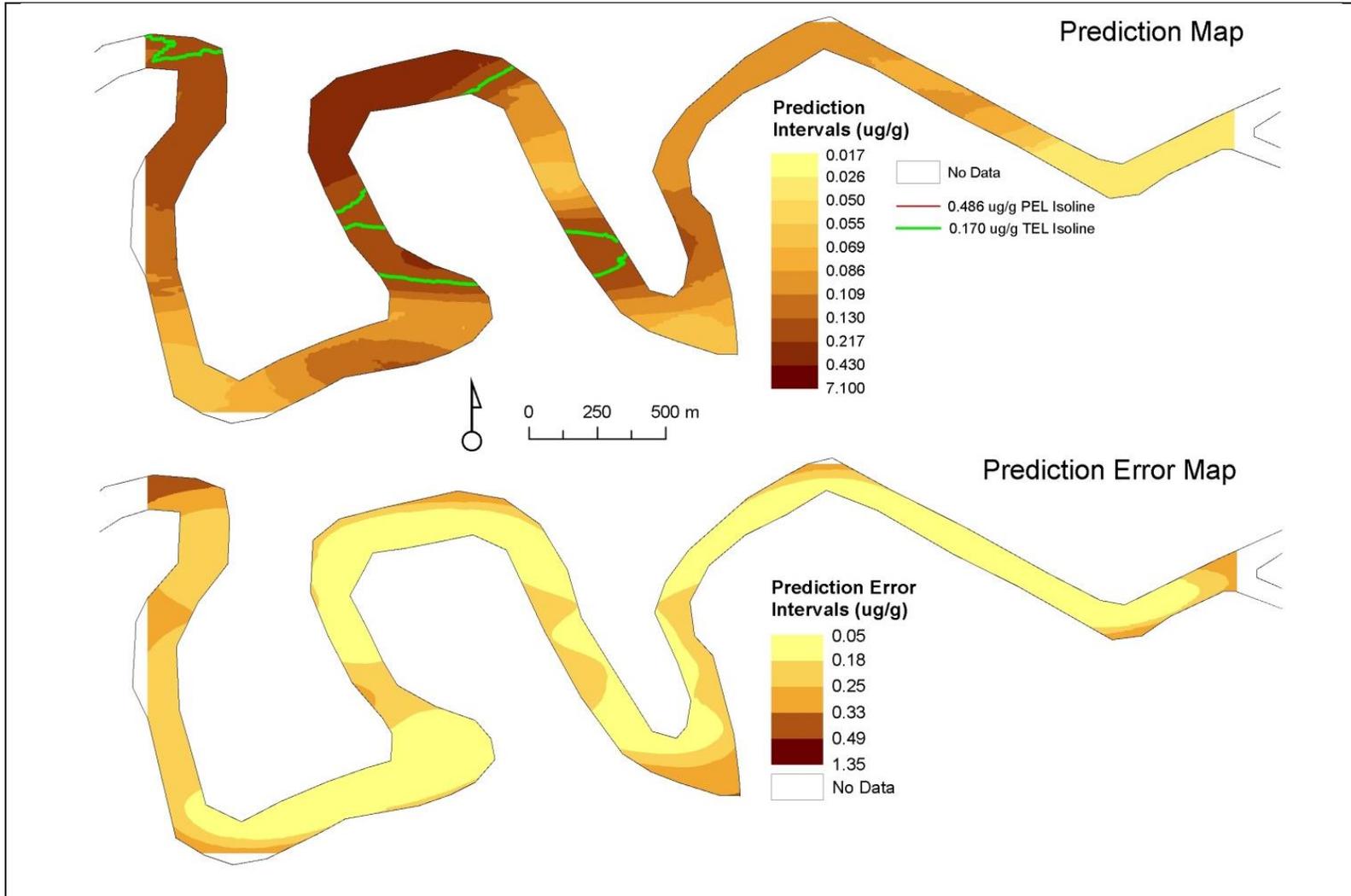
Figure 4.4 shows the predicted distribution for Hg in surface sediments and the associated prediction error map. This map was created using the 114 samples that made up the

study area data set. Hg shows a regional trend of low values in the eastern region and high values in the western region of the study area. Also, a region of very high values exists in the second north meander (NM2).

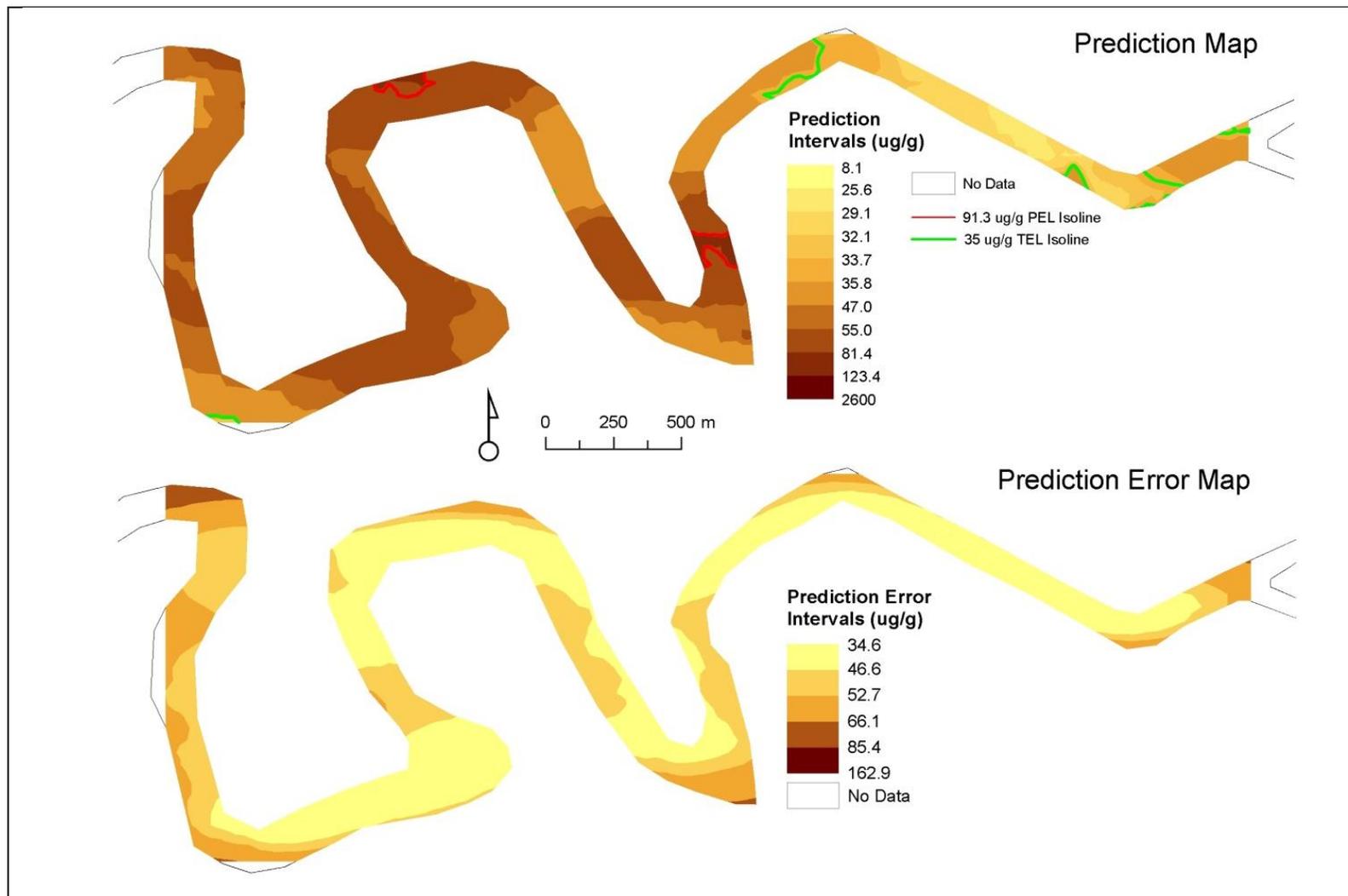
The map in Figure 4.4 also shows that no Hg PEL isolines are found within surface sediments, rather only TEL isolines exist. In fact, four different TEL regions are clearly delimited by the green TEL isolines. In terms of error, most of the map is associated with the smallest prediction error interval ( $<0.18 \mu\text{g/g}$ ). A pocket of high error values ( $>0.49\mu\text{g/g}$ ) is found in the west end of the study area.

The map of Pb's distribution, which is shown in Figure 4.5, was created using the complete dataset ( $n=114$ ). A similar trend to that of Hg can be seen, this is, there are low values of Pb in the east and high values in the west of the study area. Moreover, like the error map of Hg, that of Pb shows that the smallest prediction error class covers most of the study area and only a small proportion of area is covered with higher values.

In contrast to the distribution of Hg, the map of Pb shows three short PEL isolines and several TEL isolines. The first PEL area is found in NM2 whereas the second area is found in the second southern meander (SM2). The TEL isolines for Pb enclose a much larger region of the study area compared to those of Hg.



**Figure 4.4.** The geographical distribution of Hg in surface sediments in the study area. In the text, this map is referred as the standard Hg map.



**Figure 4.5.** The geographical distribution of Pb in surface sediments in the study area. In the text, this map is referred as the standard Hg map.

Conspicuously different Hg and Pb maps were produced when the dataset was divided into Sections A, B and C. Figure 4.6 shows the Hg distribution maps for Section A, B and C, which have been spliced together to produce a semi-continuous map.

Three major Hg PEL areas stand out in Section B (Figure 4.6), as well as two TEL areas in Section A. The TEL isolines seen in Figure 4.4 are still present in Section B; however, they show greater definition. A greater heterogeneity in the distribution of Hg can be appreciated compared to Figure 4.4.

With regards to Pb similar phenomena occur when the dataset was divided into sections. For instance in Figure 4.7, Section A shows two small PEL areas, Section B shows three larger PEL areas and Section C shows two small PEL areas, most of which are absent in Figure 4.5. Moreover, more TEL areas, both big and small, are present in all three sections.

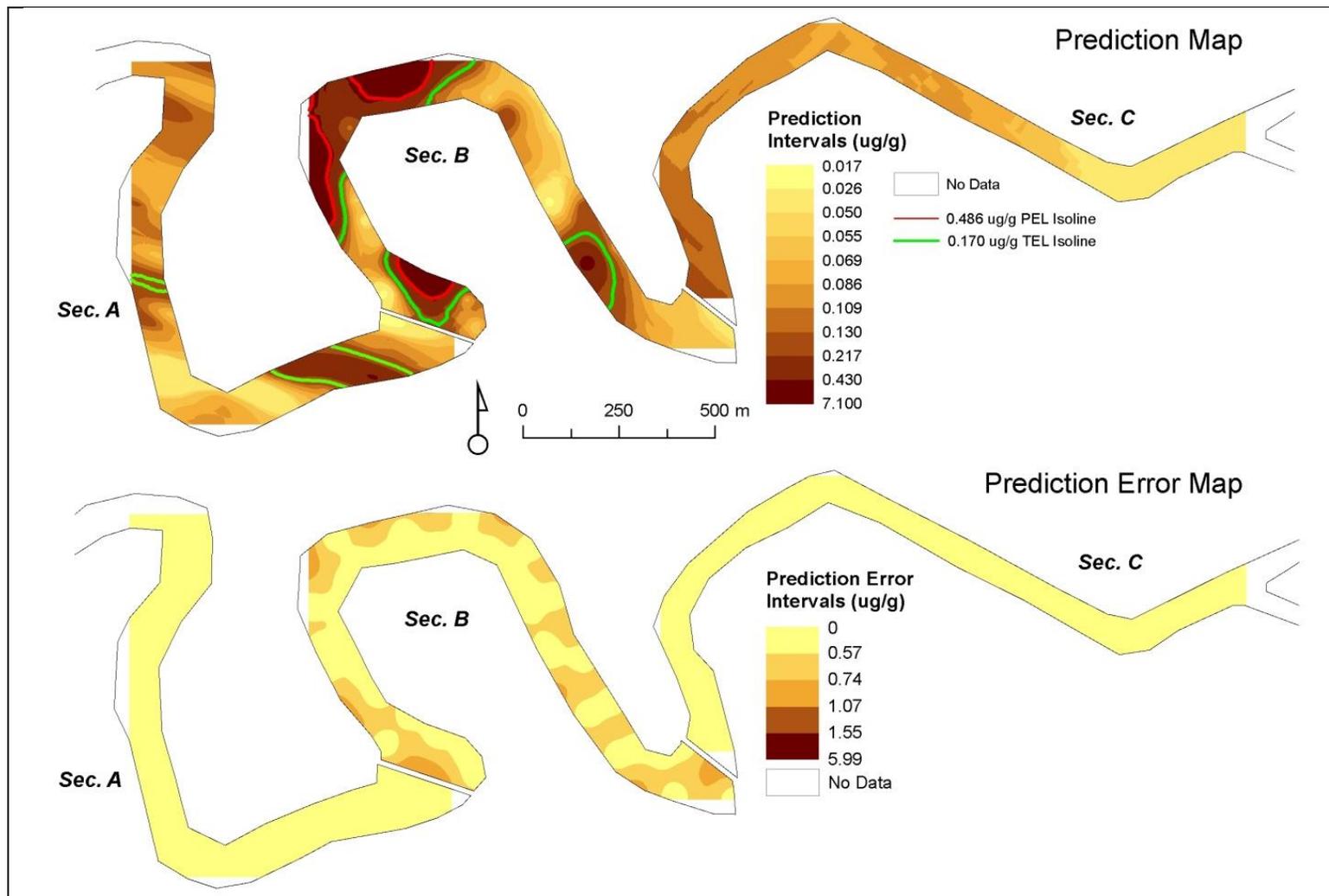
The prediction error maps of the Hg and Pb spliced maps are very similar to each other. Sections A and B in Figures 4.6 are virtually the same as those in Figure 4.7 but Section C has relatively higher values (i.e. darker tones) in the map of Pb.

It is important to note the very contrasting patterns in the eastern part of SM3 in Figures 4.6 and 4.7. This part has an Hg PEL area located in the pool (Figure 4.6) whereas the map for Pb (Figure 4.7) shows a PEL area on the opposite side of the pool, in the point bar. Moreover, the map of Hg shows that SM3 has a small PER area right in the middle whereas the map of Pb shows that most of SM3 is a PER area.

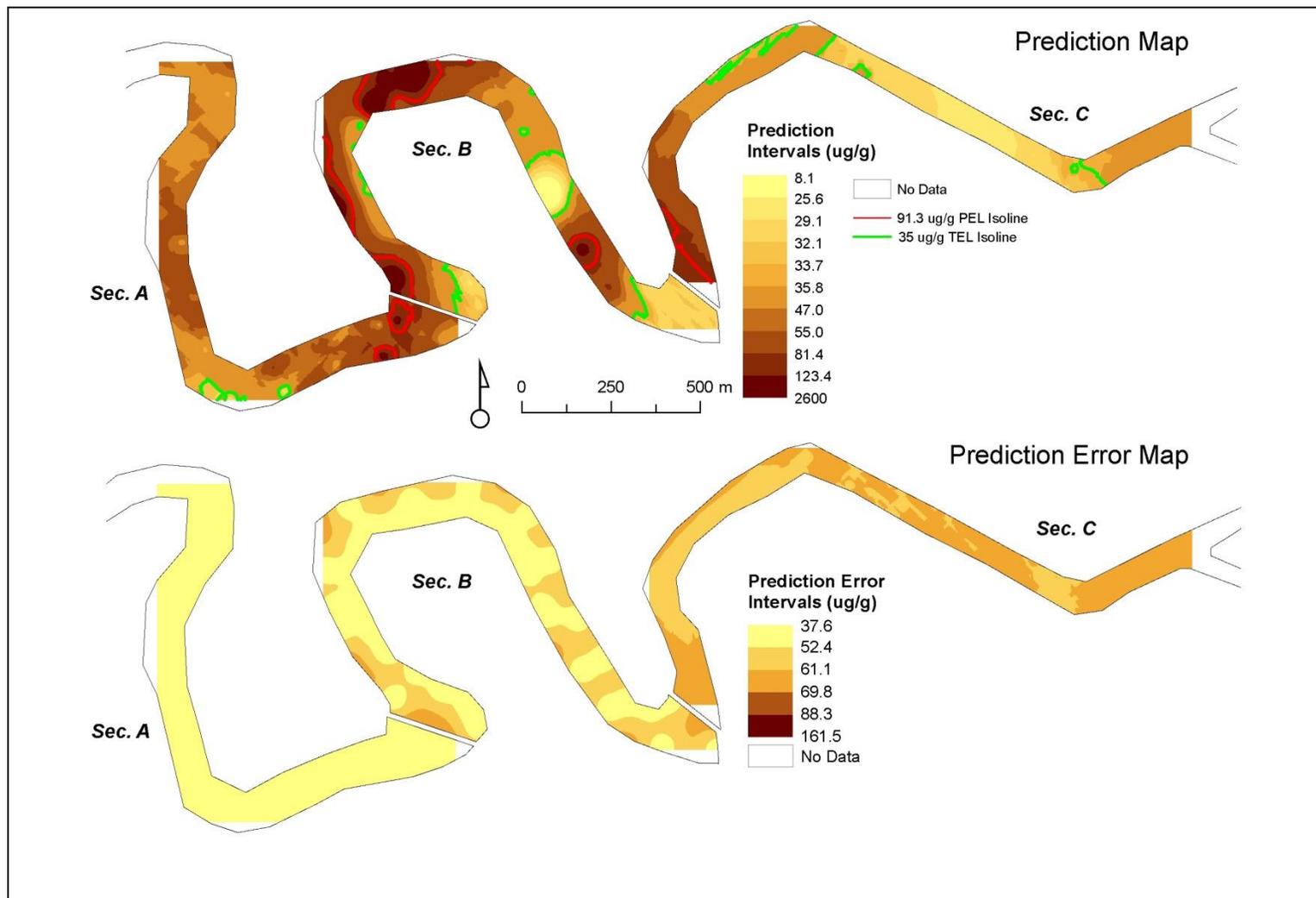
Another important characteristic worth mentioning is the mismatch in prediction values that occurs in the areas where the study area sections join. For instance, there is an abrupt change from low values to high values at the border of Section A and Section B

in both the prediction and prediction error maps for Hg and Pb in Figures 4.6 and 4.7. Similarly, there is an abrupt change at the border between Section B and Section C. Interestingly, the border between Section A and B of the prediction map for Pb (Figure 4.7) shows a smoother continuation of values compared to the map of Hg (Figure 4.6).

Finally, it is worth mentioning that the ranges of prediction error values for the standard and spliced maps were different. The Hg standard map errors range from 0.05 to 1.35  $\mu\text{g/g}$ , whereas in the spliced map they range from 0 to 5.99  $\mu\text{g/g}$ . In contrast, the Pb error range is slightly smaller in the spliced map (37.6 to 161.5  $\mu\text{g/g}$ ) than in the standard map (34.6 to 162.9  $\mu\text{g/g}$ ). These differences seem to reflect the variation with each dataset. For instance, as seen in Table 4.1 the standard deviation for the Hg study area dataset is 0.288  $\mu\text{g/g}$ , which is smaller than the average for the three sections, 0.568  $\mu\text{g/g}$ . The opposite is true for Pb, where the standard deviation of the study area dataset is larger (248.67  $\mu\text{g/g}$ ) than the average for all three sections (180.08  $\mu\text{g/g}$ ).



**Figure 4.6.** The geographical distribution of Hg in surface sediments in the study area (spliced map). The three sections were spliced together to produce this map. In the text, this map is referred as the spliced Hg map.



**Figure 4.7.** The geographical distribution of Pb in surface sediments in the study area (spliced map). The three sections were spliced together to produce this map. In the text, this map is referred as the spliced Pb map.

#### **4.4 Distribution of Hg and Pb in subsurface samples**

The maps of Hg distribution in subsurface sediments are shown in Figure 4.8. Depth 1 shows an extensive Hg PEL region (Figure 4.8A) that starts at the western edge of NM1 and ends at the western edge of NM2. Depth 1 also shows Hg TEL isolines in SM3 and on the western most riffle of the study area.

Depth 2 (Figure 4.8B) is similar to Depth 1; however, a PEL isoline shows up between the SM2 and NM2. Also, the TEL areas have contracted in SM3 and a short TEL isoline has appeared at NM1.

At Depth 3 (Figure 4.8C) the distribution of Hg is again different. The PEL isoline near NM3 has shifted southwards towards the riffle and has been replaced by TEL isolines. Similarly, new TEL isolines appear in SM2 and a more defined TEL isoline is seen in NM1. The darker PEL region found in the riffle between SM2 and NM2 has become lighter and only a small pocket of high values remains.

Depth 4 (Figure 4.8D) shows more Hg PEL isolines than the other depth classes, although the PEL areas are relatively smaller. Only a small sliver of high values is found in the riffle between SM2 and NM2. The TEL isolines present in the western end of the river at Depth 3 are absent at Depth 4. Instead, TEL isolines appear in the western edge of SM3 and an extra TEL isoline appears in the riffle between SM1 and NM1.

In summary, several general Hg concentration trends can be noted. First, the region between NM1 and NM2 has the highest concentration values within a given depth level and across depth levels. The eastern edge of SM3 together with the eastern most riffle of the study area show less pollution within any given depth and across all depths.

A part from these general patterns the TEL and PEL regions vary in size and location across all four depths.

Figure 4.9 shows the distribution of Pb in subsurface sediments at four depth levels. Depth 1 (Figure 4.9A) shows high values of Pb along the riffle between NM1 and SM2. This region is contained within an extensive PEL region that spans from the western edge of NM1 to the eastern edge of SM3. Also, at Depth 1 a smaller PEL area is found in the eastern edge of NM3 and a single TEL isoline is found in the eastern most riffle of the study area.

At Depth 2 (Figure 4.9B) the range of Pb values has decreased as seen by the absence of dark tones at this depth level. Moreover, although the larger PEL region seen in Depth 1 has remained about the same size, it has shifted eastward. Also, new PEL isolines appear in SM2. Furthermore, one PEL isoline has disappeared in NM3 effectively expanding the PEL area in this meander. No TEL isolines are found in this depth class.

In Figure 4.9C (Depth 3) the area enclosed by PEL isolines has contracted and split into two smaller regions when compared to Depth 1 and 2. Moreover, at Depth 3 two small PEL areas are seen in the western-most riffle. Furthermore, a single TEL isoline appears in the eastern end of the study area.

Finally, the Pb map for Depth 4 shows the presence of more PEL isolines. For instance, small PEL areas have appeared in SM2 and in the middle of SM3. Also, the PEL area found in the western-most riffle has expanded and now contains darker tones of brown when compared to Depths 1, 2 and 3. Moreover, two areas of high values have

appeared in the riffles between NM1 and SM2 and NM2 and SM2. Furthermore, no TEL isolines are found at Depth 4.

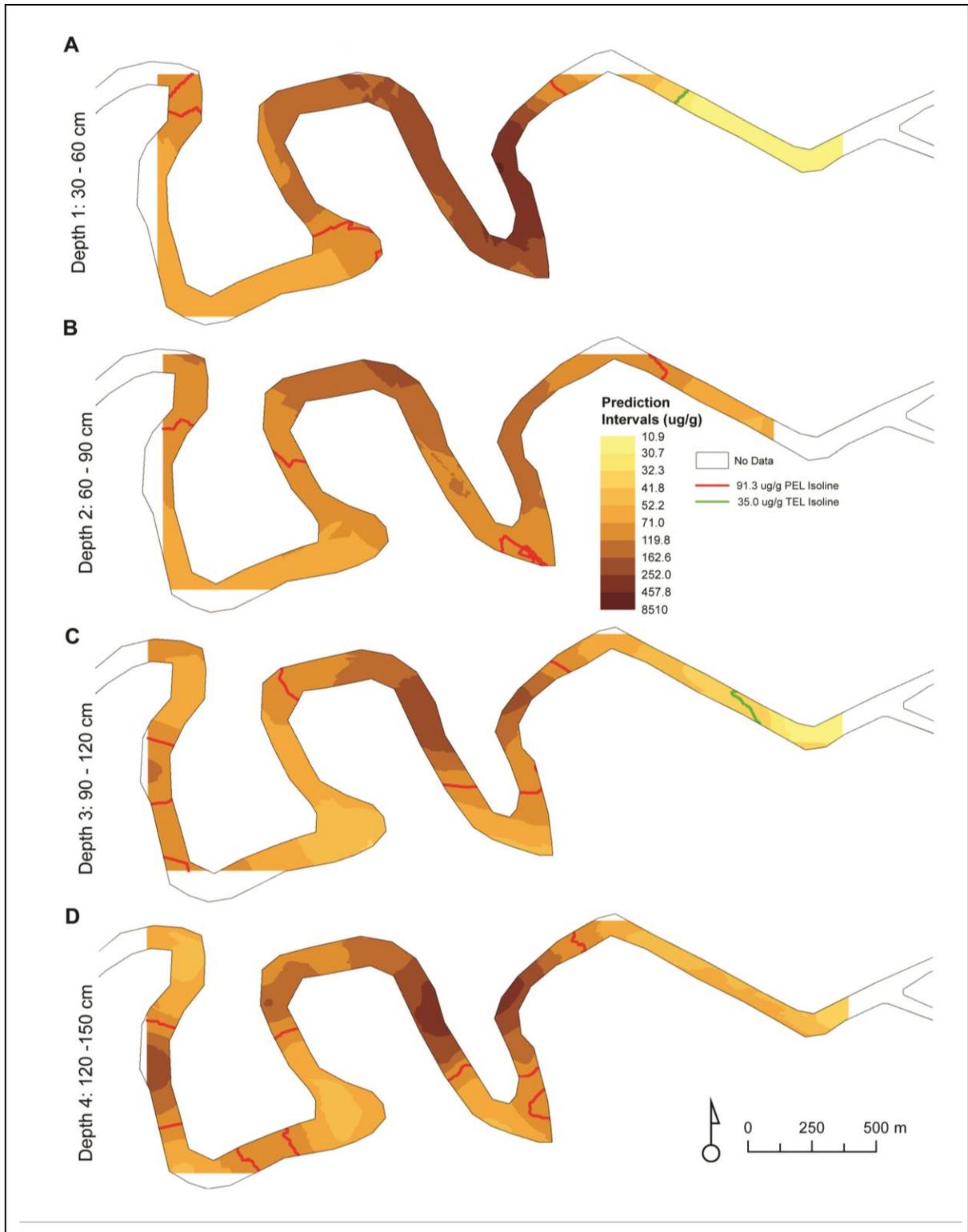
Overall it seems that Pb levels decrease with increasing depth in the region between NM1 and NM2. However, the western most region of the study area seems to show an increase in Pb levels with increasing depth. The eastern most riffle of the study area shows the lowest Pb levels across all depth levels but darker tones increase with depth.

Comparing Hg and Pb, it can be seen that the depth maps for Pb have less TEL and more PEL isolines than the subsurface sediment maps of Hg. Conversely, the region between NM2 and NM1 contains the areas with the highest concentrations of Hg and Pb across all depth levels.

The corresponding error maps for Figures 4.8 and 4.9 can be found in Appendix B. Figure B.1 shows that most of the study area is associated with relatively low Hg prediction errors in Depth 1 to 3. Only Depth 4 shows a mixture of error values throughout the map. In contrast, Figure B.2 shows Depth 1 for Pb is associated entirely with the two highest error intervals. Depth 3 shows that the lowest Pb error intervals covering most of the study area.



**Figure 4.8.** The geographical distribution of Hg in the study area at four depth levels.



**Figure 4.9.** The geographical distribution of Pb in the study area at four depth levels.

## **Chapter 5: Discussion**

### **5.1 Exploratory Data Analysis**

Histograms for Hg and Pb showed a positively skewed distribution. This shape is very common in sediment and soil contamination studies as well as in geological studies (Clark, 1979; Houlding, 2000). Even after log transformation, the distributions of Hg and Pb did not become completely normal and the values of Pb remained less normal than those of Hg. As the standard errors produced by kriging analysis are dependent on the normality of the data, it can be argued that in general terms the prediction error maps of Hg are more accurate than those of Pb.

The exploratory data analysis that was carried out illustrated that Hg and Pb showed different variability in each study area section. In terms of surface sediments, the highest Hg value and variability was found in NM2 whereas the highest Pb and variability value was found in riffle between SM2 and NM1. However, when several Hg concentration values were plotted against Pb ones, statistically significant positive correlations were seen; especially among subsurface samples (Figure 4.2). The strong positive correlation in older (deeper) sediments might indicate common sources of pollution.

A basic assumption of kriging is that the difference in values at two points is related to the distance and direction between these two points and nothing else (Clark, 1979). In other words, kriging assigns weights to each sampled point to predict unknown values based in two criteria, similarity in terms of distance (and direction) and autocorrelation of data values (Houlding, 2000). These criteria are not quite valid in the case of a meandering river such as the one studied in this paper. The Euclidean distance

used in kriging does not take into account the dry land that is found between meanders. As a consequence, the weights assigned to sample points found within the semi-variogram range are not proportional to the real in-river distance.

## **5.2 Semi-variograms and Kriging analysis**

The use of the ArcGIS default parameters acted as a stepping stone for the exploration and uncovering of optimal kriging parameters. For instance, the theoretical semi-variogram model that scored the best using the default parameters remained the best semi-variogram for all other kriging models for a given data set. This meant that after the initial three kriging models that used default parameters, the best theoretical semi-variogram was determined and used as the sole semi-variogram in all other remaining kriging models.

In this paper, only Hg concentrations for Section C were best modeled using the ArcGIS default values. In all other cases, non-default parameters provided more accurate models. However, it seems that the gains in accuracy obtained by using more involved models do not translate to substantial improvements in areal maps. This is specially noted in the prediction error and Q-Q plots of the best and worst models for a given data set, which did not show striking differences.

A deeper examination of semi-variogram elements can reveal important characteristics of the pollutants being modeled. Figure 5.1 shows that the sill and nugget for the best kriging models are positively correlated. A big nugget was associated with a big sill in virtually all cases, except for Hg in Section B (highest point on y-axis of Figure 5.1). This positive correlation is not surprising as the nugget contributes to the value of

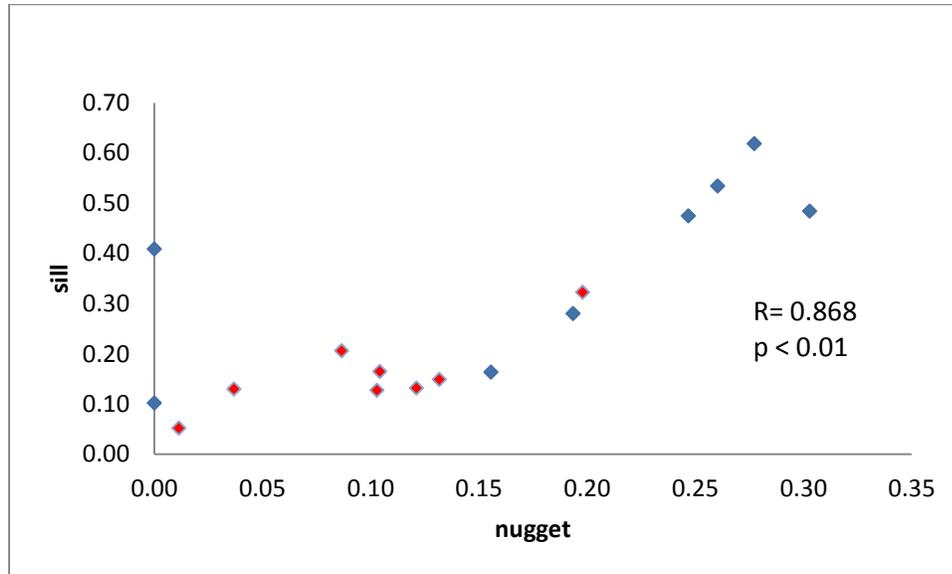
the sill. What is surprising is that the nugget of Hg was very close to zero in Sections A and B. This means that the semi-variogram did account for most of the local Hg variation in these sections. In the case of Section B, this is counterintuitive as the likelihood of inaccurately weighting neighbouring sample points in this section is high, given that it has the most convoluted meanders of all the study area.

As mentioned in Chapter 2, the nugget accounts for sampling errors and micro-scale variation (i.e. randomness) that cannot be modeled by the semi-variogram and the sill is equivalent to the variance of a data set (Clark, 1979). Table A.1 in Appendix A shows that the nuggets for Pb in Sections A and B are bigger than those for Hg, whereas in Section C the opposite is the case. Interestingly though, the nuggets for both Hg and Pb in Section C are much larger than their counterparts in Sections A and B. Consequently, the distribution of Pb in the surface sediments of Section A and B is more random than that of Hg but more predictable than that of Hg in Section C. Additionally, the distribution of both metals is less predictable in Section C compared to the other two sections.

The relatively high nuggets of Hg and Pb in the study area data sets compared to those in the study area section data sets, support the idea that prediction surfaces created with the former data sets contain a much larger random component in their estimated values. Only Pb for Section C does not conform to this notion due to its larger nugget size compared to Pb for the study area data set.

In subsurface sediments Hg variability also seemed to be affected by significant micro-scale variation and sampling error. This was reflected on the larger nugget of Hg compared to that of Pb. In fact, within subsurface samples, the average nugget for Hg

(1.1) was about twice the average nugget for Pb (0.52). For this reason, it may be argued that the distribution of Hg in subsurface sediments is less predictable than that of Pb.



**Figure 5.1.** The relationship between sill and nugget for all data sets. The values for Hg (blue) and Pb (red) are plotted together.

It is important to note that core loss as well as diagenetic effects (i.e., physical or chemical processes that affect sediments after they have been deposited) will contribute toward an increased nugget effect (Clark, 1979). This is especially true for Hg, which has a high propensity to change its concentration due to natural physical and chemical processes (D'Itri and D'Itri, 1977). It has been estimated that the contribution of sampling error to the nugget is only about 3% (Clark, 1979). Thus, it can be said that most of the nugget is the result of micro-scale or random variation.

The parameters that had the strongest influence on the accuracy of kriging models were lag size and number of lags. As a rule of thumb it has been suggested that the lag size times the number of lags should be less than half the largest distance between any two points in the data set (Johnston *et al.*, 2001; Wackernagel, 2003). As the largest

distance between any two points in the study area dataset was 2,907 m, the rule of thumb for the study area translates to about 1,450 m.

In this paper, the best combination of lag size and number of lags for Hg and Pb was 76 m and ten respectively. Multiplying these two numbers together a total distance of 760 m was obtained, which is about half the rule of thumb distance. This combination of lag size and lag number was obtained directly and indirectly from Ouyang *et al.* (2002; 2003a). In the case of the Cedar-Ortega river system, Ouyang *et al.* (2003a) found that for Hg the best combination was a lag size of 160 m and the number of lags equal to ten. In the case of Pb, Ouyang *et al.* (2002) found that a lag size of 140 m was best together with ten lags. As is the case with this paper, the lag sizes found by Ouyang *et al.* (2002; 2003a) are about 26% the length of the study area. This is surprising as the shape of the Cedar and Ortega Rivers has more branches and less pronounced meanders than the Buffalo River. However, the sampling scheme in both cases was similar as they were both based on USEPA protocols. It could be argued that a similar lag size proportion and numbers of lags proved the best due to similar distances between sample points. This finding reinforced the notion that lag size and number of lags were the most crucial parameters in these kriging analyses.

In the case of subsurface sediments, the lag size-lag number product varied from 805 to 1,064 m for Hg and from 532 to 760 m in the case of Pb. This showed that the bulk of variability of Hg and Pb in surface sediments was found within similar areal extents, whereas in deeper sediments this variability was found in larger areas for Hg and in smaller areas for Pb. It is important to note that, although the extents for surface and

subsurface sediment data sets were similar, the sample sizes were smaller for the latter data sets.

Ouyang *et al.* (2002) found it necessary to vary parameters (in terms of lag size, nugget effect, sill and range) in their analyses of different pollutants. In the present paper, except for lag size and number of lags, no set of parameters proved to be the best for a given data set or type of metal analyzed. Whereas Ouyang *et al.* (2002, 2003a, 2003b) only tested spherical isotropic semi-variograms, this paper tested three different theoretical semi-variograms with and without the incorporation of directional influences. Contrary to Ouyang *et al.* (2002, 2003a, 2003b), this paper found that Gaussian and exponential models proved best for most of the data sets while anisotropy was relevant in six out of sixteen best kriging models.

The present research illustrated that accounting for real world phenomena does not necessarily produce the most accurate kriging model (i.e. highest BMI value). For instance, this paper tested polynomial trends in order to account for the flow of the Buffalo River. Moreover, the search neighbourhood was converted to an ellipse with its major axis overlaid on the path of the river. This was performed in order to give more weight to in-river neighbouring sample points. Despite the efforts to account for the actual physical nature of the river in the modeling process, the resulting kriging models were sometimes among the least accurate models produced. The question then arises as to whether to choose the model with higher accuracy or that which accounts for known geographical patterns.

### **5.3 Standard and Spliced prediction maps**

The Buffalo River presents special challenges and opportunities to study the geographical distribution of Hg and Pb. One challenge arises from the fact that it is a meandering river and therefore, kriging analysis may produce biased results if improper weighting of sample points occurs.

Inaccurate weighting is likely as the Geostatistical Analyst offers limited control over the search neighbourhood and this cannot bend to conform to the shape of the study area. Thus, two points which are on different meanders are grouped as neighbours by the software although there is land between them. To mitigate this issue, two mapping approaches were taken to create prediction maps for the concentration of Hg and Pb within surface sediments.

The standard maps for Hg (Figure 4.4) and Pb (Figure 4.5) are substantially different than the corresponding spliced maps (Figure 4.6 and 4.7). The first set of maps show a continuous prediction surface whereas the second sets of maps show a gap and abrupt jumps in values at the splice joints. More importantly, the first set of maps only show global trends whereas the spliced maps show regional and local trends.

Standard and spliced maps are useful in their own ways. For instance, it could be argued that for restoration and sediment removal purposes the spliced maps are better as they more accurately pin point the regions of high pollution. Conversely, standard maps can be produced relatively easily and still capture global trends. It should be noted that the distinction between standard and spliced maps is only justifiable when dealing with irregularly shaped study areas. In all other cases standard maps will not only be the most accurate but also will reveal any local, regional or global trends.

Both types of maps cannot be compared in terms of prediction errors because to do so requires the same extent and sample size in each prediction surface being compared. This was clearly not the case as the standard maps covered the entire study area; whereas the spliced maps were created using three different surface maps, each covering about a third of the study area. Moreover, each standard map was created using 114 sample points; whereas the spliced maps were produced with datasets that contained 33 to 41 sample points each.

It would have been instructive to test other distance estimation methods such as the coordinate transformation (Barabas *et al.*, 2001) or in-river distance network (Gardner *et al.*, 2003). Both of these methods take into consideration the meandering nature of rivers to improve prediction surfaces but without the need to split the study area and data set as has been done in this paper. Although novel, both of these methods require more intricate data processing such as the creation of matrices to transpose coordinates (Barabas *et al.*, 2001) and the calculation of as-the-fish-swims distances via network analysis (Gardner *et al.*, 2003). Unfortunately, these methods were beyond the scope of the present study.

Insight into why standard and spliced maps are different from one another can be gained by comparing the standard prediction map of Hg (Figure 4.4) with that of Hg in Section A (Figure 4.6). It is clear that during the kriging process the sample points located in the NM2 contributed more of the weights than sample points located up or down stream. This is to be expected due to the fact that the NM2 is closer in Euclidean distance to the western end of the study area. However, this does not make sense geographically as there is dry land between these two regions. More weight should have

been given to the sampling points which are closer in terms of in-river distance. If this would have been the case, then the interpolated values in the western end of the study area would have been as low as those seen in NM3 of Section A.

Some authors have compared the prediction results obtained using Euclidean distance and in-river distance with mixed outcomes. Some studies have shown that Euclidean distance is a good enough measure (Little *et al.*, 1997), whereas other studies have shown that in-river distance substantially improves predictions (Gardner *et al.*, 2003). This paper supports the latter notion because dividing the study area in sections that are less convoluted has “tricked” the kriging algorithm into giving more weight to in-river neighbouring points, which became the closest in Euclidean distance as well.

The double approach used in this paper allowed for the comparison and contrast of results obtained using global (standard maps) and regional data sets (spliced maps). It should be noted that this technique would have not been needed if a less convoluted river system was studied.

#### **5.4 Geographical distribution of Hg and Pb in the Buffalo River**

The distribution of Hg and Pb in surface and subsurface sediments of the study area is both similar and different in several ways. For instance, Figure 4.1 shows that the concentrations of Hg and Pb in subsurface sediments are on average much higher than those in surface sediments. In contrast, by examining Figures 4.1A and 4.1B individually, it can be seen that the concentration of Hg increases with depth, whereas that of Pb decreases with depth.

The sedimentation patterns found in this paper are different than those found in other studies. For instance, Ouyang *et al.* (2002, 2003a) and Fabbri and Trevisani (2005), found that the concentrations of Hg and Pb were highest in surficial sediments and decrease with depth. This discrepancy could probably be explained by the dredging that took place in the river two years before the samples were collected. Dredging could have removed surficial sediments with higher levels of Hg and Pb concentrations than those found in this study. Alternatively, it is also possible that by the time sampling took place the inputs to the river system were already reduced.

The kriging surfaces obtained in this study show a global trend in Hg and Pb concentrations which consists of relatively low values in the east and high values in west of the study area. This pattern of deposition is common in rivers that have an association with industrial activity. Metal concentrations tend to increase downstream in rivers found in industrial areas as the river flows from undeveloped to developed parts of the watershed (Miller and Orbock-Miller, 2007). An interesting follow up study would be to include the location of past and present, actual and potential, points sources and see how close or far are these from the areas of high pollution.

The study area has a particular sediment pattern that is different from other rivers. The mid section of the study area (i.e. area between SM2 and NM2) has higher values than the far west region. In fact, the spliced maps for Hg and Pb show that Section B has more PEL and TEL hotspots than Section A.

It seems that the hotspots of high concentration values seen in the spliced maps reflect the pollution caused by point sources for three main reasons. First, most of the Hg and Pb hotspots in surface sediments are located closer to the river banks rather than the

mid region of the river. Second, Hg and Pb seem to follow a typical contaminant transport and dispersion pattern with high concentrations near the source and lower concentrations away from the source (Ouyang *et al.*, 2003). Thirdly, point source pollution is more likely to have happened than diffuse source pollution because the study area has been the home of several industries, many of which were located right next to the river (Sutton, 2006).

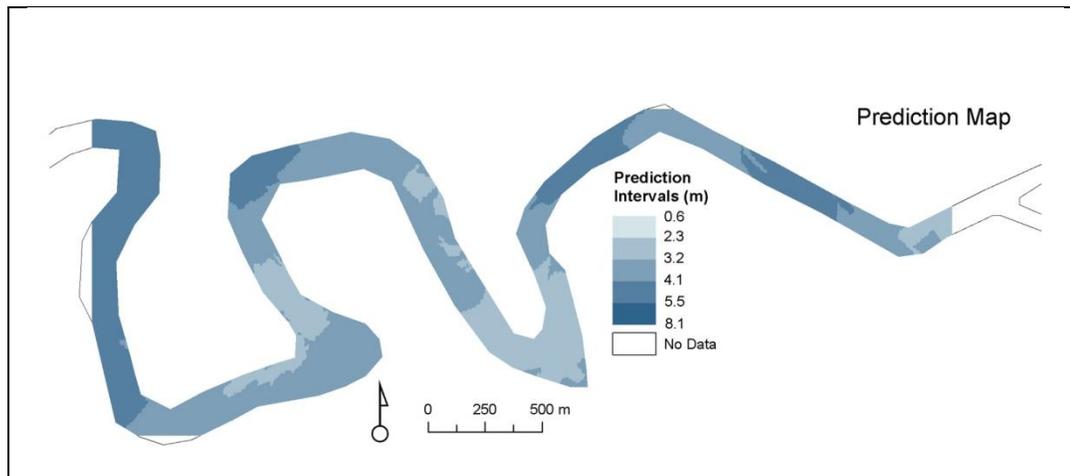
A noteworthy sediment pattern is that the shape of the PEL and TEL isolines in the spliced Hg and Pb prediction maps are more elongated in the direction of the river flow than otherwise. These contours could reflect the influence of water flow in shaping the distribution of sediments and thus, the distribution of Hg and Pb concentrations.

This paper shows that most of the meanders are associated with high concentrations of Hg and Pb. Meanders tend to accumulate not only sediments but also within these a greater proportion of fine grain particles, which are known to play a strong role in pollutant binding (Miller and Orbock Miller, 2007). The spliced maps for Hg (Figure 4.6) and Pb (Figure 4.7) show that most of the PEL isolines are found within or near meanders. SM2 is an exception as it shows some of the lowest Hg and Pb concentration values of the study area, second only to the eastern most riffle.

Another interesting association worth exploring is that of pollutant concentration and water depth. Ideally, a bathymetric map would have been better to explore this relationship more accurately; however, the data available only allowed for the creation of a depth prediction map. Figure 5.2 shows a prediction map of water depth values along the study area. Here it can be seen that the riverbed has a mixture of deep and shallow areas. Moreover, it can be seen that most of the meanders have mid range depth values

except for small areas within NM2 and NM3, which have the high depth values. Interestingly, the riffle between NM3 and SM3 has a uniform high depth value range (5.5 – 8.1 m).

If the surface sediment maps were analyzed together with the water depth map, then the following observations could be made: 1) unlike other studies where a strong association between water depth and high values of Hg and Pb concentration was found (Forsythe *et al.*, 2004), high pollutant concentrations in this study area were found in both shallow and deep water regions. Some of the highest values of Hg and Pb are found in areas with mid-depth values and some of the lowest concentration values are found within high depth values. For instance, the riffle between NM2 and SM3 which has mid-depth values also has two hotspots of high Hg and Pb pollution.



**Figure 5.2.** Prediction map of water depth along the study area.

The standard maps of Hg and Pb show more similarities than dissimilarities. For instance, the general trend of low values in the east and higher values in the west is evident in both maps. Also, both maps show high concentration regions in the meanders

especially in NM2. In terms of dissimilarities, the Hg standard map only shows TEL isolines whereas the Pb standard map shows both PEL and TEL isolines.

The spliced maps highlight more differences in the sedimentation patterns of Hg and Pb. For example, the distribution of high Pb values is patchier than that of Hg. Although the map of Pb (Figure 4.7) shows more TEL and PEL areas, their sizes are rather small. In contrast, the Hg prediction map (Figure 4.6) has fewer TEL and PEL areas but these are larger in extent. These sedimentation pattern differences indicate that the study area has a more pronounced Hg pollution that covers a relatively small area and less intense Pb pollution that covers a relatively large area.

The more extensive and less intense pollution of Pb compared to that of Hg is more clearly seen in subsurface sediment maps (Figures 4.8 and 4.9). Pb has only two TEL isolines at all four depth levels whereas the maps for Hg show several TEL isolines at each depth level. Looking at each depth level individually, it can be seen that Depth1 has the highest Pb values of all depths. The highest Pb values are mostly found in the areas between the western edge of SM2 and NM2 across all depths. Depth 2 and 3 show pockets of high values in the western most riffle of the study area. Similarly, Hg also shows the highest values in the areas between the western edge of SM2 and NM2 across all depths and PEL pockets in the western most riffle.

The positive correlation of Hg and Pb values is not only seen in the maps but also clearly shown in correlation plots of log Pb against log Hg (Figure 4.2). As illustrated in the plots there is a strong positive correlation among subsurface sediments. Conversely, surface sediment samples show a much weaker positive correlation. Other studies have also found a positive correlation between Hg and Pb in surficial sediments (Norville,

2005). The weak correlations seen in surface sediments in the study area might also be the result of dredging. The positive correlation seen in the study area's sediments might reflect a hidden but stronger correlation between the amount of fine grain sediments and metal concentrations.

The control samples taken by NYSDEC were located upstream of the study area. The values found for Hg and Pb, were 0.016  $\mu\text{g/g}$  and 5.8  $\mu\text{g/g}$  respectively. Comparing these control values to the prediction maps values, it can be seen that the former are out of the range of values presented in the maps. The lowest values found within the study area are 0.017  $\mu\text{g/g}$  for Hg and 8.1  $\mu\text{g/g}$  for Pb. The heavy pollution within the study area is such that even the six bottom most samples had much higher Hg and Pb concentration (0.13  $\mu\text{g/g}$  and 39.4  $\mu\text{g/g}$  respectively) than the control samples. On average, Hg concentration in the study area was about seven times higher than the control and Pb was about five times higher.

In summary, the distributions of Hg and Pb along the river share similarities and differences at both the surface and subsurface levels. These characteristics are not only seen in the prediction maps but also in the kriging parameters, histograms, and correlation plots. Although more quantitative analysis is needed, the results of this paper suggest that the best remedial work for the study area might be not to remove top layers of sediment because as this paper has shown, subsurface sediments are more polluted than surface ones. Removing the top layer may expose pockets of high Hg and Pb pollution which may act as new sources of contamination for the study area. If remedial work is to be undertaken then the spliced maps might provide a good reference for the selective removal of only Hg and Pb PEL hotspots.

## 5.5 Study Limitations and Future Work

This paper has several issues and limitations that are worth mentioning. For instance, log-transformation has a tendency to underestimate values and their probabilities because too much weight is given to low values (Clark, 1979; Houlding, 2000). In order to overcome this scale distortion, the use of a correction factor has been suggested by Houlding (2000). This correction is given by,

$$\text{Correction factor} = \text{antilog}(\text{estimated values}) \times \frac{\text{avg. original values}}{\text{avg. estimated values}} \quad (4)$$

The impact of log transformation on predicted values was tested using the Hg data set for the entire study area. It was seen that the prediction values increased by about 7% compared to non-corrected ones. Perhaps, a full implementation of the correction factor may be warranted in order to improve the accuracy of the predictions as an increase in values is likely to translate into more TEL and PEL regions.

A second limitation is that ordinary kriging assumes an unknown but constant mean; that is, the mean is not a function of location only the error (Johnston *et al.*, 2001). This assumption may not be valid for the study area due to its pronounced meandering nature. Nonetheless, the flexibility of ordinary kriging is likely to produce relatively accurate predictions.

Another limitation of this study is that the effect of periodic dredging within the study area sediments was not fully analyzed. It is obvious that such a disturbance would

substantially alter the distribution of Hg and Pb. The ways in which dredging alters sediment pollution distributions would be a valid research topic.

An issue to keep in mind is that core sampling only provides averages, which means that there is a loss of information (Miller and Orbock Miller, 2007). Moreover, core sampling does not provide information on the range of pollutant concentrations that exist within the core. Thus, the prediction maps presented in this paper may under predict the real number and extent of TEL and PEL isolines.

It is important to note that spatial patterns vary in time (Miller and Orbock Miller, 2007) and the data used in this study were collected through a synoptic sampling scheme (i.e. samples collected from a large number of locations over a short period of time) (Sutton, 2006). Consequently, the present assessment provides only a snap shot of past Hg and Pb pollution conditions within the study area. The present distribution of these metals is likely to be different; especially, if it is considered that the Buffalo River's watershed yields about 86,719 tons of sediment per year (Shreeram, 2004).

Finally, in terms of future work it might prove valuable to explore the usefulness of co-kriging analysis using water depth or percent clay as the second variable. The extra information contributed by these variables might improve the accuracy of the prediction maps. Furthermore, it would be very informative to carry out ordinary kriging analysis using as-the-fish-swims distance rather than as-the-crow-flies distance. As shown in this paper, irregularly shaped study areas need special consideration in order to properly calculate the distances among sample points.

## **Chapter 6: Conclusions**

The patterns of deposition for Hg and Pb found in this paper show that their spatial distribution is variable and different. In general terms both of these pollutants show a trend toward decreasing concentrations going from east to west of the study area. This trend is not linear as the mid region of the study area shows higher pollution than the western most region.

Pb pollution in the study area was shown to be more widespread but less pronounced than Hg pollution. Pb had many more PEL regions than Hg in both surface and sub-surface sediments. However, Hg had PEL regions with relatively higher concentrations than those of Pb.

Documenting the geographical distribution of Hg and Pb sheds light on the location of major areas of pollution and possibly into the sources of pollution. To achieve the latter, more detailed studies are required in terms of mapping the location of past and present sources of potential pollution along the banks of the Buffalo River.

The prediction maps presented in this paper together with in-depth chemical and physical analysis of Hg and Pb will help in understanding the underlying processes that lead to the predicted patterns. An important factor in discerning these depositional patterns will be to consider the effect of periodic dredging in the Buffalo River.

The study area provided a great opportunity to explore the kriging process itself and how to go about creating prediction surfaces for irregularly shaped study areas. This paper showed the value of creating regional maps as these highlight important regional and local variations, which can be overlooked if only global maps are used. Furthermore,

this paper sheds light on the fact that deep dredging may cause renewed pollution as it will expose deeper and more contaminated sediments.

Any plans for remedial and restorative work in the Buffalo River Area of Concern should focus on the PEL hotspots identified in the prediction maps. These areas are relatively small and could be cleaned up in a more cost-effective manner.

## References

Barabas, N., Goovaerts, P. and Adriaens, P. 2001. Geostatistical assessment and validation of uncertainty for three-dimensional dioxin data from sediments in an estuarine river. *Environmental Science and Technology*. 35, 3294-3301.

Bookstrom, A.A., Box, S.E., Campbell, J.K., Foster, K.I. and Jackson, B.L. 2001. Lead-rich sediments, Coeur d'Alene River valley, Idaho: area, volume, tonnage, and lead content. *USGS Publications*. 1-79.

Botkin, D.B. and Keller, E.A. 2005. *Environmental Science: earth as a living planet* (5th Edition ed.). Hoboken: John Wiley and Sons, Inc.

Canadian Council of Ministers of the Environment. 2001. *Canadian sediment quality guidelines for the protection of aquatic life: Introduction*. In: *Canadian Environmental Quality Guidelines, 1999*. Winnipeg: Canadian Council of Ministers of the Environment.

Cheung, K.C. and Wong, M.H. 2006. Risk assessment of heavy metal contamination in shrimp farming in Mai Po Nature Reserve, Hong Kong. *Environmental Geochemistry and Health*. 28(1-2), 27-36.

Clark, I. 1979. *Practical Geostatistics*. Essex: Applied Science Publishers.

D'Itri, P.A. and D'Itri, F.M. 1977. *Mercury Contamination: a human tragedy*. New York: John Wiley and Sons, Inc.

Environmental Systems Research Institute Canada Ltd. (ESRI). *ArcCanada 3.0 – Continental Data – North America: Water* [digital resource: vector]. Version 3.0. Toronto, Canada: Environmental Systems Research Institute Canada Ltd., 2003. Retrieved on April 15, 2009, from <https://www.runner.ryerson.ca/madar/geospatial/libdata/>

Environmental Systems Research Institute (ESRI). 2001, August. *ArcGIS Geostatistical Analyst: statistical tools for data exploration, modelling and advanced surface generation*. Retrieved 05 30, 2009, from ESRI White Papers: <http://www.esri.com/library/whitepapers/pdfs/geostat.pdf>

Fabbri, P. and Trevisani, S. 2005. A geostatistical approach to a pollution case in Northeastern Italy. *Mathematical Geology*. 37 (6), 569-586.

Forsyth, J. 2005, March 16. *Mercury Pollution, Autism Link Found - U.S. Study*. Retrieved April 18, 2009, from CommonDreams.org: <http://www.commondreams.org/headlines05/0317-05.htm>

- Forsythe, K.W. and Marvin, C.H. 2009. Assessing historical versus contemporary mercury and lead contamination in Lake Huron sediments. *Aquatic Ecosystem Health and Management*. 12 (1), 101-109.
- Forsythe, K.W., Dennis, M. and Marvin, C.H. 2004. Comparison of Mercury and Lead Sediment Concentrations in Lake Ontario (1968–1998) and Lake Erie (1971–1997/98) using a GIS-Based Kriging Approach. *Water Quality and Resources Journal of Canada*. 39 (3), 190-206.
- Gardner, B., Sullivan, P. J. and Lembo, A. J. 2003. Predicting stream temperatures: geostatistical model comparison using alternative distance metrics. *Canadian Journal of Fisheries and Aquatic Science*. 60, 344-351.
- Gonzalez-Weller, D., Karlsson, L., Caballer, A., Hernandez, F., Gutierrez, A., Gonzalez-Iglesias, T., Marino, M. and Hardisson, A. 2006. Lead and cadmium in meat and meat products consumed by the population in Tenerife Island, Spain. *Food Additives and Contaminants*. 23(8), 757-763.
- Harrison, R.M. and Laxen, D.P. 1981. *Lead Pollution, Causes and Control*. London: Chapman and Hall Ltd.
- Houlding, S.H. 2000. *Practical Geostatistics - Modeling and Spatial Analysis*. Berlin: Springer.
- Johnston, K., Ver Hoef, J.M., Krivoruchko, K. and Lucas, N. 2001. *Using ArcGIS Geostatistical Analyst*. Redlands: ESRI.
- Lewis, J. 1985, May. *Lead Poisoning - An Historical Perspective*. Retrieved May 16, 2009, from U.S. Environmental Protection Agency: <http://www.epa.gov/history/topics/perspect/lead.htm>
- Little, L.S., Edwards, D. and Porter, D.E. 1997. Kriging in estuaries: as the crow flies, or as the fish swims? *Journal of Experimental Marine Biology and Ecology*. 213, 1-11.
- Miller, J.R. and Orbock Miller, S.M. 2007. *Contaminated Rivers - A geomorphological-geochemical approach to site assessment and remediation*. Dordrecht: Springer.
- Norville, W. 2005. Spatial distribution of heavy metals in sediments from the Gulf of Paria, Trinidad. *Revista de biología tropical*. 53, 33-40.
- Ouyang, Y., Higman, J., Campbell, D. and Davis, J. 2003. Three-dimesnisonal kriging analysis of sediment mercury distribution: a case study. *Journal of the American water Resources Association*. 689-702.

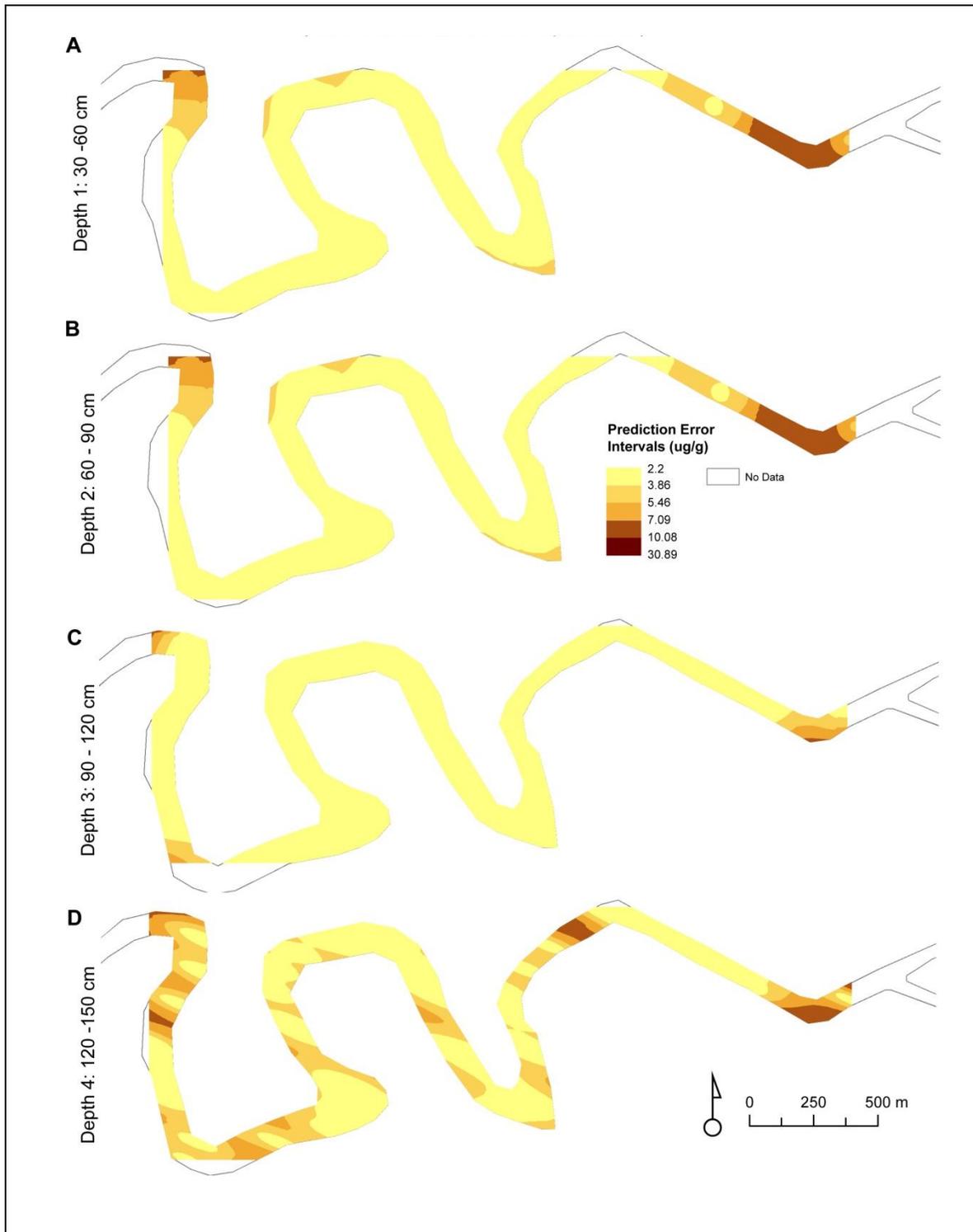
- Ouyang, Y., Higman, J., Thompson, J., O'Toole, T. and Campbell, D. 2002. Characterization and spatial distribution of heavy metals in sediment from Cedar and Ortega rivers subbasin. *Journal of Contaminant Hydrology*. 54, 19-35.
- Ouyang, Y., Nkedi-Kizza, P., Mansell, R.S. and Ren, J.Y. 2003. Spatial Distribution of DDT in Sediments from Estuarine Rivers of Central Florida. *Journal of Environmental Quality*. (32), 1710-1716.
- Palmer, R. F., Blanchard, S., Stein, Z., Mandell, D. and Miller, C. 2006. Environmental mercury release, special education rates, and autism disorder: an ecological study of Texas. *Health and Place*. 12 (2), 203-209.
- Shi, J., Wang, H., Xu, J., Wu, J., Liu, X. and Zhu, H. 2007. Spatial distribution of heavy metals in soils: a case study of Changxing, China. *Environmental Geology*. 52, 1-10.
- Shreeram, I. 2004. *Sediment Modelling for the Buffalo River Watershed*.
- Smol, J. P. 2008. *Pollution of Lakes and Rivers - A paleoenvironmental perspective* (2nd Edition ed.). Malden: Blackwell Publishing.
- Sutton, G.P. 2006. *Buffalo Sediment Study*. New York State Department of Environmental Conservation, Buffalo.
- Time. 1966, September 23. *Lead Among Romans*. Retrieved May 16, 2009, from Time Magazine: <http://www.time.com/time/magazine/article/0,9171,842832,00.html>
- University of Wisconsin-Madison. 2006, August 11. *Mercury Pollution Threatens Health Worldwide, Scientists Say*. Retrieved April 18, 2009, from ScienceDaily: <http://www.sciencedaily.com/releases/2006/08/060811191845.htm>
- United States Environmental Protection Agency (US EPA ). 2008, May 2008. *Buffalo River Area of Concern*. Retrieved May 22, 2009, from US Environmental Protection Agency: <http://www.epa.gov/glnpo/aoc/buffalo.html>
- Wackernagel, H. 2003. *Multivariate Geostatistics - an introduction with applications*. Berlin: Springer.

## Appendix A

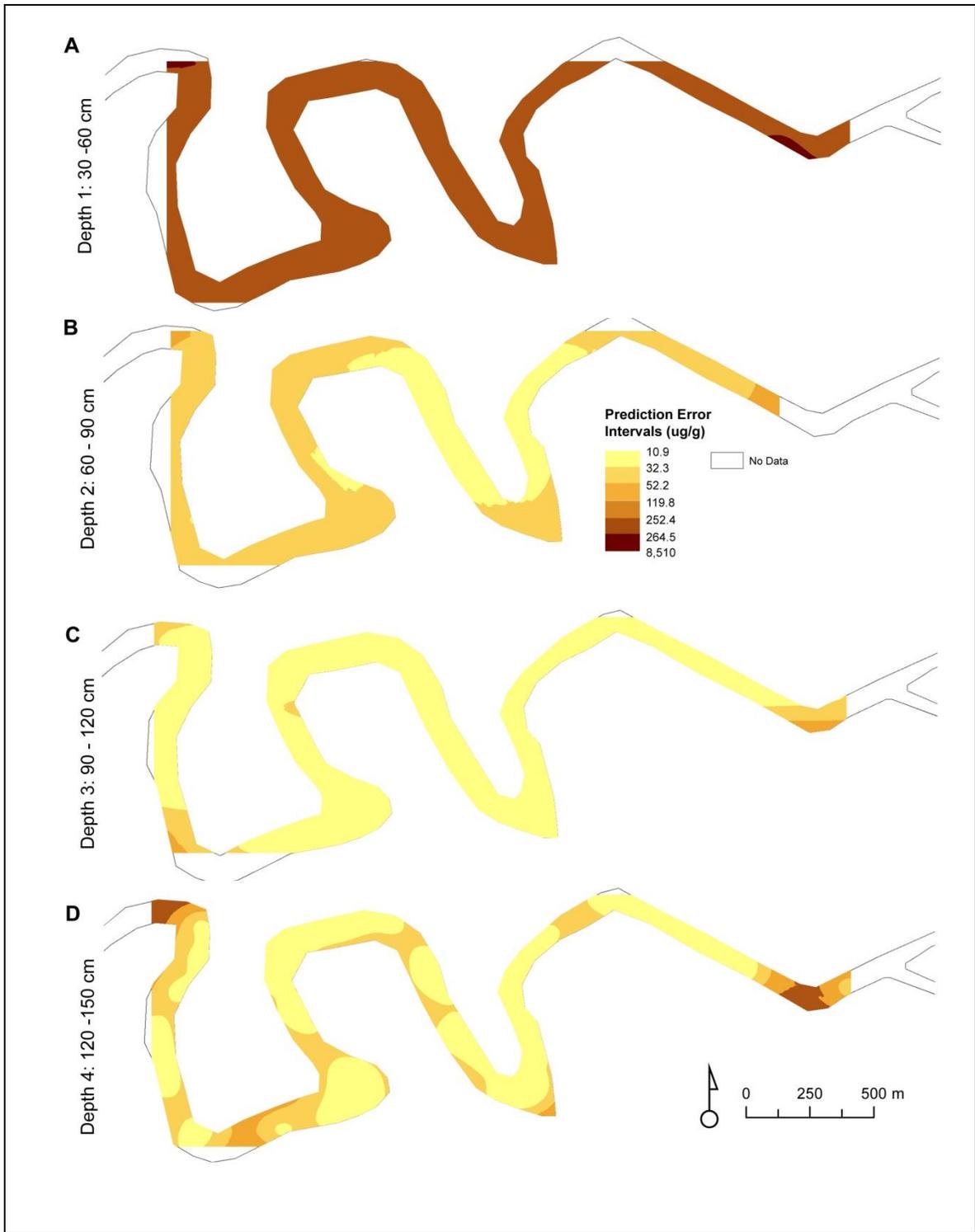
**Table A.1.** The best kriging models and their parameters. The models shown below are those that scored the highest using the best model index (BMI). An ellipse with four sectors was used for all models and the number of neighbours was set to 5 with the inclusion of at least two neighbours. The largest distance between two points was 2,900 m. (G, Gaussian; S, spherical; Ex, Exponential; Rg., range; S-A, semi-axis;  $\gamma$  theoretical semi-variogram; n, sample size.) Units of nugget, partial sill and sill are  $\mu\text{g}/\text{g}^2$ . Units of ranges, search axes and lag size are in metres.

Var.	n	Trend	$\gamma$	Ani.	Nugget	Partial Sill	Sill	Maj. Rg.	Min. Rg.	Lag Size	No. Lags	Angle	Maj. S-A	Min. S-A	M	RMS	ASE	MS	RMSS	BMI
Study Area (full data set)																				
Hg	114	no	G	yes	0.193	0.087	0.280	754.3	246.1	76	10	275	754.3	246.1	0.000	0.482	0.482	-0.001	1.003	2.995
Pb	114	no	G	yes	0.103	0.025	0.128	754.3	241.3	76	10	290	1290	600	0.001	0.345	0.345	0.002	1.002	2.996
Section A																				
Hg	33	no	S	yes	0.000	0.102	0.102	268.5	93.6	22.6	12	288	268.5	93.6	0.001	0.293	0.287	0.006	1.006	2.982
Pb	33	no	Ex	no	0.011	0.041	0.052	77.4	null	39	12	0	77.4	77.4	-0.008	0.249	0.242	-0.032	1.03	2.932
Section B																				
Hg	40	2nd	Ex	no	0.000	0.409	0.409	211.4	null	28.6	10	0	211.4	211.4	-0.018	0.566	0.567	-0.018	0.981	2.962
Pb	40	1st	S	no	0.037	0.093	0.130	207.9	null	17.5	12	321	800	190	-0.007	0.313	0.322	-0.003	0.982	2.970
Section C																				
Hg	41	no	S	no	0.156	0.008	0.164	496.2		50	10	0	496.2	496.2	-0.013	0.428	0.417	-0.030	1.023	2.937
Pb	41	no	Ex	no	0.121	0.011	0.132	694.7	null	70	10	0	694.7	694.7	0.000	0.388	0.371	0.000	1.039	2.946
Depth1																				
Hg	34	2nd	Ex	no	0.303	0.181	0.484	800.9	null	115	7	0	800.9	800.9	0.004	0.654	0.670	0.002	0.991	2.974
Pb	34	no	S	no	0.198	0.125	0.323	713.1	null	76	10	0	713.1	713.1	0.006	0.518	0.536	0.008	0.996	2.971
Depth2																				
Hg	33	no	G	no	0.277	0.341	0.619	867.1	null	76	14	0	867.1	867.1	0.002	0.610	0.617	-0.001	1	2.992
Pb	33	1st	G	no	0.132	0.018	0.149	754.3	null	76	10	0	754.3	754.3	-0.002	0.386	0.390	-0.008	0.995	2.983
Depth3																				
Hg	49	1st	S	no	0.247	0.228	0.475	497.9	null	115	7	0	497.9	400	0.001	0.623	0.634	0.004	1	2.985
Pb	49	1st	G	yes	0.104	0.061	0.165	754.3	281.8	76	10	294	1290	600	0.002	0.376	0.376	0.006	1.002	2.992
Depth4																				
Hg	50	1st	Ex	yes	0.260	0.274	0.534	900.8	232.7	76	12	294	1290.9	600	-0.006	0.693	0.697	0.005	1.001	2.990
Pb	50	no	G	yes	0.087	0.120	0.206	529.28	440.2	76	7	273.7	1290	600	0.003	0.395	0.398	0.007	1.006	2.984

## Appendix B



**Figure B.1.** Prediction error maps for Hg in the study area at four depth levels.



**Figure B.2.** Prediction error maps for Pb in the study area at four depth levels.