

VISUALIZING THE *BIG THREE*: GEOSPATIAL INTERPOLATION OF HEAVY METAL
SEDIMENT CONTAMINATION IN LAKE ERIE

By

Danielle Elizabeth Mitchell

Master of Spatial Analysis, Ryerson University, 2014

Bachelor of Arts (Honours), University of Ottawa, 2013

A dissertation

presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in the Program of

Environmental Applied Science and Management

Toronto, Ontario, Canada, 2018

© Danielle Elizabeth Mitchell, 2018

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A DISSERTATION

I hereby declare that I am the sole author of this dissertation. This is a true copy of the dissertation, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this dissertation to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this dissertation 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.

I understand that my dissertation may be made electronically available to the public.

ABSTRACT

Visualizing the *Big Three*: Geospatial Interpolation of Heavy Metal Sediment Contamination in Lake Erie

Danielle E. Mitchell

Doctor of Philosophy in the Program of Environmental Applied Science and Management, 2018

Ryerson University

A wealth of resources for economic prosperity have driven development along the shorelines of the Great Lakes for over 150 years. The rapid growth of industrial, agricultural, and residential land use has degraded many natural components of lake ecosystems, including sediments and water quality. In this dissertation, spatiotemporal patterns of non-essential heavy metal sediment contamination in Lake Erie will be examined from historic and contemporary sediment surveys. Three inter-related studies explore innovative methods for improving the validity and overall usefulness of sediment contamination maps that could be used by a variety of stakeholders in pollution control efforts throughout the Lake Erie basin. First, sediment survey designs are analyzed for their utility in creating valid interpolated surfaces from which spatiotemporal comparisons of mercury sediment contamination can be compared over time. The next study explores how ancillary sediment variables and contamination categorization methods can support interpolated maps of cadmium sediment contamination from low-density sediment surveys. The final study introduces a novel method of three-dimensional geovisualization to enhance the geographic representation of lead sediment contamination patterns throughout the Lake Erie basin. Innovative research methodologies designed for this dissertation may be applied to sediment contamination studies in other Great Lakes. The visualization techniques employed in mapping sediment contamination patterns provide strong scientific evidence for spatiotemporal change in non-essential heavy metal pollution throughout Lake Erie. Combined, the research findings and maps produced throughout this dissertation can contribute to the growing body of knowledge used in environmental decision making for pollution control in the Great Lakes basin.

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisor Dr. K. Wayne Forsythe for guiding me through to the timely completion of this dissertation. Over four years and two degrees, Dr. Forsythe has encouraged me to be an honest, fearless, and upstanding individual for which I am truly grateful. Thank you to the members of the examination committee and to Dr. Chris Marvin and Debbie Burniston from Environment and Climate Change Canada for your scrutiny and constructive feedback. I would also like to thank my dear friend Dr. Christopher Greene, the Mitchell, Harrington, and Ford families for their unwavering faith and moral support.

*To Cam, Jacob, Cait,
and the girl who pulled it all off.*

TABLE OF CONTENTS

Author's Declaration	ii
Abstract	iii
Acknowledgements	iv
Dedication	v
List of Tables	ix
List of Figures	xi
List of Important Acronyms and Abbreviations	xiv
Chapter 1: Introduction	1
1.1 Research Context	1
1.2 Defining Contamination	2
1.2.1 Mercury (Hg)	2
1.2.2 Cadmium (Cd)	3
1.2.3 Lead (Pb)	3
1.3 The Practicality of Spatial Interpolation	4
1.4 Focus on Lake Erie	6
1.5 Research Objectives and Goals	7
1.6 Chapter Summaries and Authorship Statements	8
1.6.1 Chapter 2: Assessing Statistical and Spatial Validity of Sediment Survey Design and Sampling Densities: Examples from Lake Erie	8
1.6.2 Chapter 3: Temporal Trends and Origins of Lake Erie Cadmium Contamination in Relation to Sediment Substrate Type Using Multivariate Kriging Analyses	9
1.6.3 Chapter 4: Multivariable 3D Geovisualization of Historic and Contemporary Lead Sediment Contamination in Lake Erie	9
1.6.4 Chapter 5 – Contributions to Great Lakes Research and Future Directions	9
Chapter 2: Assessing Statistical and Spatial Validity of Sediment Survey Design and Sampling Densities: Examples from Lake Erie	10
2.1 Introduction	11
2.1.1 Data Collection	12
2.1.2 Geovisualization of Sediment Contamination	12
2.1.3 Kriging	12
2.2 Data	15
2.3 Methods	15
2.3.1 Down-sampled Survey Density	17

2.3.2 Density Change Detection.....	18
2.3.3 Subset Kriging Analysis	20
2.3.4 Categorical Change Detection Analysis	27
2.4 Discussion	31
2.5 Conclusion	33
Chapter 3: Temporal Trends and Origins of Lake Erie Cadmium Contamination in Relation to Sediment Substrate Type Using Multivariate Kriging Analyses	34
3.1 Introduction	35
3.1.1 Fate of Cd Sediment Contamination in Aquatic Environments	35
3.1.2 Geochemical and Geomorphological Associations to Cd Contamination	36
3.2 Study Area and Data	36
3.2.1 Cd Contamination in Lake Erie	37
3.2.2 Sediment Substrate Classification.....	37
3.3 Methods	39
3.3.1 Index of Geoaccumulation (I_{geo})	39
3.3.2 Canadian Sediment Quality Guidelines for the Protection of Aquatic Life (CCME)	40
3.3.3 Co-Kriging	41
3.3.4 Kriging	41
3.3.5 Change Detection	42
3.3.6 Statistical Analysis	43
3.3.6.1 1971 Regression	43
3.3.6.2 1997/1998 Regression	44
3.4 Spatial Analysis	45
3.5 Raster Overlay and Sediment Substrate Analysis	50
3.6 Discussion and Conclusion	52
Chapter 4: Multivariable 3D Geovisualization of Historic and Contemporary Lead Sediment Contamination in Lake Erie.....	56
4.1 Introduction	57
4.1.1 Pb Contamination in Lake Erie	58
4.1.2 Lake Erie Circulation	59
4.1.3 Geovisualization	60
4.2 Data	61
4.2.1 Kriging	61

4.2.2 Canadian Sediment Quality Guidelines for the Protection of Aquatic Life (CCME)	62
4.3 Methods	62
4.3.1 Kriging	62
4.3.2 3D Geovisualization	64
4.3.3 Change Detection	66
4.4 Spatial Analysis	66
4.4.1 Kriging Results	66
4.4.2 Spatiotemporal Change in Pb Sediment Contamination	69
4.5 Discussion	73
4.5.1 Historic and Contemporary Pb Sediment Contamination	73
4.5.2 Multivariate Geovisualization	74
4.6 Conclusion	77
Chapter 5: Conclusion	78
5.1 Status of the <i>Big Three</i> in Lake Erie	78
5.2 Validity of Spatially Interpolated Contamination Maps	79
5.3 Avenues of Future Research Applications	80
References	82

LIST OF TABLES

Table 1-1: Cross-validation statistics, with conventional standards of accuracy (ArcGIS for Desktop, 2016)	5
Table 2-1: Descriptive statistics of Hg sediment contamination samples from Lake Erie	15
Table 2-2: Descriptive statistics of Hg sediment contamination from the complete and down-sampled subsets Acknowledgements	18
Table 2-3: Kriging cross-validation statistics of the complete dataset and all down-sampled subsets.....	19
Table 2-4: Hg contamination intervals, and respective thresholds, used for the density change detection analysis	19
Table 2-5: Calculated contamination change categories and their meanings	20
Table 2-6: Spatial extent of predicted Hg contamination by subset	23
Table 3-1: Descriptive statistics of Cd sediment contamination samples from Lake Erie	37
Table 3-2: The Wentworth (1922) Scale of Grain Sizes (Williams et al., 2011)	39
Table 3-3: Index of geoaccumulation and pollution classifications (Müller, 1981).	40
Table 3-4: Kriging parameters used for modeling the 1971, and 1997/1998 datasets	42
Table 3-5: Kriging cross-validation statistics of the 1971 and 1997/1998 sediment surveys	42
Table 3-6: Cd values organized by whole contamination intervals	43
Table 3-7: Calculated contamination change categories and their meanings	43
Table 3-8: Correlations between Cd values from 1971 and auxiliary variables	44
Table 3-9: Correlations between Cd values from 1997/1998 and auxiliary variables.....	44
Table 3-10: Surface area (km ²) coverage of Cd contamination ≥TEL to <PEL and ≥PEL by substrate type from the 1971 dataset	51
Table 3-11: Surface area (km ²) coverage of Cd contamination <TEL and ≥TEL to <PEL by substrate type from the 1971 dataset	52
Table 4-1: Descriptive statistics of Pb sediment contamination samples from Lake Erie	61
Table 4-2: Kriging cross-validation statistics of the 1971, 1997/1998, and 2014 Pb datasets ...	63

Table 4-3: Kriging cross-validation statistics of the 1971, 1997/1998, and 2014 Pb (log-normal) datasets.....	63
Table 4-4: Kriging parameters used for modeling the 1971, 1997/1998, and 2014 Pb (log-normal) datasets	63
Table 4-5: Pb (log-normal) contamination intervals <TEL, ≥TEL to <PEL, and ≥PEL.....	66
Table 4-6: Contamination change categories and their meanings.....	66
Table 4-7: Spatial extent of predicted Pb (log-normal) contamination by year	67
Table 4-8: Surface area change in Pb (log-normal) sediment contamination between 1971 and 1997/1998	70
Table 4-9: Surface area change in Pb (log-normal) sediment contamination between 1997/1998 and 2014	71
Table 4-10: Surface area change in Pb (log-normal) sediment contamination between 1971 and 2014.....	72

LIST OF FIGURES

Figure 1-1: Example semivariogram displaying the fit of a Gaussian model through the 1971 Hg data points.....	5
Figure 1-2: The Emerging Great Lakes Megaregion encompasses the entire Lake Erie shoreline, extending inland to several major American cities, and municipal regions of southwestern Ontario (RPA, 2008)	6
Figure 1-3: Organizational structure of dissertation research	8
Figure 2-1: Lake Erie and its basins (Marvin et al., 2002; Esri, 2017)	11
Figure 2-2: Categorical dot map of Hg sediment contamination from surveys of Lake Erie conducted in a) 1971 and b) 2014.....	13
Figure 2-3: Annual Lake Erie circulation patterns adapted from Beletsky et al., 1999; the two-gyre system in the central basin is identified individually by the a) anticyclonic gyre and b) cyclonic gyre.....	16
Figure 2-4: Point density maps of a) 90%, b) 80%, c) 70%, d) 60%, e) 50%, f) 40%, g) 30%, h) 20%, and i) 10% randomly down-sampled subsets from the complete dataset.....	17
Figure 2-5: Kriging predictions of Hg sediment contamination in Lake Erie based on the complete dataset	21
Figure 2-6: Kriging predictions of Hg sediment contamination in Lake Erie based on a) 90%, b) 80%, and c) 70% subsets of the complete dataset.....	24
Figure 2-7: Kriging predictions of Hg sediment contamination in Lake Erie based on d) 60%, e) 50%, and f) 40% subsets of the complete dataset	25
Figure 2-8: Kriging predictions of Hg sediment contamination in Lake Erie based on g) 30%, h) 20%, and i) 10% subsets of the complete dataset.....	26
Figure 2-9: Change detection maps of no change, over-, and underestimations of Hg sediment contamination from kriging predictions based on a) 90%, b) 80%, and c) 70% subsets of the complete dataset	28
Figure 2-10: Change detection maps of no change, over-, and underestimations of Hg sediment contamination from kriging predictions based on 60% (d), 50% (e), and 40% (f) subsets of the complete dataset	29
Figure 2-11: Change detection maps of no change, over-, and underestimations of Hg sediment contamination from kriging predictions based on 30% (g), 20% (h), and 10% (i) subsets of the complete dataset	30
Figure 2-12: Kriging predictions of Hg sediment contamination in Lake Erie based on the 2014 dataset	31

Figure 3-1: The Emerging Great Lakes Megaregion encompasses the entire Lake Erie shoreline, extending inland to several major American cities, and municipal regions of southwestern Ontario (RPA, 2008)	37
Figure 3-2: Lake Erie substrates, re-drawn from (Haltuch et al., 2000)	38
Figure 3-3: Kriging results of the 1971 Cd contamination in Lake Erie	45
Figure 3-4: Kriging results of the 1997/1998 Cd contamination in Lake Erie	46
Figure 3-5: Change in Cd contamination concentrations and extent between 1971 and 1997/1998	47
Figure 3-6: Anthropogenic pollution intensity measured by the index of geoaccumulation in contrast to the PEL isoline for the 1971 Cd dataset	49
Figure 3-7: Anthropogenic pollution intensity measured by the index of geoaccumulation in contrast to the PEL isoline for the 1997/1998 Cd dataset	49
Figure 3-8: Cd contamination \geq TEL to $<$ PEL and \geq PEL from the 1971 dataset by substrate type	50
Figure 3-9: Cd contamination $<$ TEL and \geq TEL to $<$ PEL from the 1997/1998 dataset by substrate type	52
Figure 3-10: Extent of anthropogenic contamination identified by the I_{geo} Index around the tributaries of Swan Creek and the Huron River in the western basin of Lake Erie. I_{geo} class = 3, 1971, outlined in dark purple; I_{geo} class = 1, 1997/1998, outlined in light purple (Esri, 2017) ...	54
Figure 4-1: Lake Erie and its basins (Marvin et al., 2002; ESRI, 2017)	57
Figure 4-2: Lake Erie depth (m) map (NOAA, 2014)	58
Figure 4-3: Annual Lake Erie circulation patterns with 20m and 50m bathymetry isolines (Beletsky et al., 1999); the two-gyre system in the central basin is identified individually by the a) anticyclonic gyre and b) cyclonic gyre	60
Figure 4-4: Lake Erie depth profile (m)	64
Figure 4-5: Bathymetry of Lake Erie viewed from the normal southern position	65
Figure 4-6: Bathymetry of Lake Erie from the southeast perspective	65
Figure 4-7: Kriging results of the 1971 Pb (log-normal) contamination in Lake Erie with average annual circulation patterns.....	67
Figure 4-8: Kriging results of the 1997/1998 Pb (log-normal) contamination in Lake Erie with average annual circulation patterns	68
Figure 4-9: Kriging results of the 2014 Pb (log-normal) contamination in Lake Erie with average annual circulation patterns.....	69

Figure 4-10: Change in Pb (log-normal) sediment contamination between 1971 and 1997/1998 70

Figure 4-11: Change in Pb (log-normal) sediment contamination between 1997/1998 and 2014 71

Figure 4-12: Change in Pb (log-normal) sediment contamination between 2017 and 2014 72

Figure 4-13: Kriging results of the 1971 Pb (log-normal) contamination in Lake Erie 75

Figure 4-14: Kriging results of the 1997/1998 Pb (log-normal) contamination in Lake Erie 75

Figure 4-15: Kriging results of the 2014 Pb (log-normal) contamination in Lake Erie 76

LIST OF IMPORTANT ACRONYMS AND ABBREVIATIONS

µg/g	Microgram per gram
2D	Two Dimensional
3D	Three Dimensional
ASE	Average Standard Error
CCME	Canadian Council of Ministers of the Environment
Cd	Cadmium
CDF	Confined Disposal Facility
DEM	Digital Elevation Model
EDA	Exploratory Data Analysis
GLIN	Great Lakes Information Network
GLWQA	Great Lakes Water Quality Agreement
Hg	Mercury
I_{geo}	Index of Geoaccumulation
IJC	International Joint Commissions
MCDM	Multiple-Criteria Decision Making
MPE	Mean Prediction Error
NCEI	National Centers for Environmental Information
NOAA	National Oceanic and Atmospheric Association
PEL	Probable Effect Level
Pb	Lead
RMSPE	Root-Mean-Squared Prediction Error
SRMSPE	Standardized Root-Mean-Squared Prediction Error
TEL	Threshold Effect Level
USAED	United States Army Corps of Engineers
USEPA	United States Environmental Protection Agency
VIF	Variance Inflation Factor

1.1 Research Context

If a picture is worth a thousand words, how about a map? A successful geographic information system (GIS) is one which can accurately convey complex, often multivariable or interdisciplinary data. Generalization is an inevitable dilemma facing all cartographers; creating maps that are scientifically sound, statistically valid and accessible to a variety of users is an arduous task. Each of the three research objectives presented in this dissertation directly address the role of scientific communication to expert and layperson knowledge production. Dr. Anne Roe stresses that “nothing in science has any value to society if it is not communicated” (Roe, 1953); this dissertation attempts to improve upon this notion by focusing on how well science can be communicated to citizens of the Great Lakes basin. For the following maps to be worth a thousand words, it is my responsibility as a geographic information scientist to ensure that valid knowledge can be derived from accurate, concise, and honest geographic visualizations.

Concern over the health and environmental sustainability of the Great Lakes is driven by a wide range of stakeholders. The Great Lakes are a vital freshwater resource for life in Canada and the United States of America. Along their combined 17,549 km shoreline (GLIN, 2017), the Great Lakes have supported centuries of economic development and prosperity in both countries, but not without consequential environmental impacts. Environmental scientists, policy makers, government agencies, academics, industry leaders, public interest groups and private citizens have supported on-going pollution control efforts since the enactment of the Canada - United States Great Lakes Water Quality Agreement (GLWQA) in 1972.

Environment and Climate Change Canada (ECCC) scientists have surveyed and archived sediment samples collected throughout the Great Lakes since 1965 (Marvin et al., 2002; Rukavina et al., 2013). Polluted sediments may be used as a proxy indicator of water quality (Forsythe et al., 2016a) and the overall health of aquatic ecosystems (Adams and Stadler-Salt, 2016). Since 2002, a group of scientists and academic researchers have mapped spatially interpolated sediment contamination values throughout many Great Lakes and tributaries (Forsythe et al., 2004; Forsythe et al., 2005; Forsythe and Marvin, 2009; Forsythe et al., 2010; Forsythe et al., 2013; Forsythe et al., 2015; Forsythe et al., 2016a; Forsythe et al., 2016b; Forsythe et al., 2016c; Gawedzki and Forsythe, 2012). The main objective of their research was to generate continuous data layers from sediment sample point data using the kriging spatial interpolation method (Forsythe et al., 2016a). Some studies (Forsythe et al., 2016a-c) also utilized bathymetric data

(recently made available by the National Oceanic and Atmospheric Administration (NOAA)) to create three-dimensional (3D) geovisualizations of sediment contamination patterns. These innovative maps offer a novel perspective over traditional two-dimensional (2D) pollution surfaces by evoking natural human perceptions of elevation and depth with colour, contrast, and illumination tools (Rinner, 2007; Lloyd and Dykes, 2011; Knust and Buchroithner, 2014).

1.2 Defining Contamination

Point (direct) and nonpoint (indirect) sources of organic and inorganic pollution enter the Great Lake basin from ad-hoc urbanization, agricultural, and industrial practices (Bogue, 2000). When absorbed in sediments, lake-wide circulation patterns result in the transport of contaminated materials throughout the lake. Sediment contamination is measured on a scale of thresholds established by the Canadian Council of Ministers of the Environment (CCME, 1999). Threshold values are unique to different organic and inorganic pollutants. Below the Threshold Effect Level (TEL), contaminated sediments are not likely to threaten human health and aquatic species. Contamination above the Probable Effect Level (PEL) is likely to pose health risks to humans and aquatic life.

Lead (Pb), Cadmium (Cd), and Mercury (Hg) are the first, fifth and seventh most predominant heavy metals in contaminated environments, respectively (Wuana and Okieimen, 2011). Known as the Big Three, they are toxic heavy metals with no known biological function (Wuana and Okieimen, 2011; Dartmouth, 2012). They are unlike Chromium (Cr), Zinc (Zn), and Copper (Cu) which are essential heavy metals commonly found in contaminated environments (Serafim et al., 2012). These heavy metals, although dangerous (and non-essential) for human development are naturally occurring in the earth's crust (and subsequently sediments in the Great Lakes), and have particular affinities for different sediment grain sizes and substrate types (ATSDR, 2007; Wuana and Okieimen, 2011; Dartmouth, 2012). In aquatic environments, elemental Hg may transform to methylmercury, a quickly bioaccumulated toxin throughout aquatic food chains (USGS, 2014). Overtime, Cd, and Pb contaminated sediments are likely to sink deeper into the sediment column, or be redistributed throughout the lake basin (Forsythe et al., 2016c).

1.2.1 Mercury (Hg)

Coal-fired power plants are the largest contributor of mercury contamination in the Great Lakes (Wuana and Okieimen, 2011, Stamper et al., 2012). Lake Erie is particularly polluted due to its proximity to Ohio, and Michigan; the two highest mercury emitting states into the Great Lakes (Stamper et al., 2012). Mercury is also released through industrial, agricultural and

pharmaceutical waste, and the combustion of fossil fuels. Concentrations in fish have been an environmental and health concern since the 1970's, where both the Canadian and American governments released fishing bans or advisories (Kot-Wasik, 2005, Stamper et al., 2012). Mercury poisoning is well known to cause neurological damage, and in some cases can lead to insanity (Kot-Wasik, 2005). Kidney damage and cancer are also long-term health concerns resulting from mercury contamination (Kot-Wasik, 2005; Wuana and Okieimen, 2011).

1.2.2 Cadmium

Cadmium pollution is a byproduct of petroleum production, electroplating, and battery manufacturing (Kot-Wasik, 2005; Wuana and Okieimen, 2011). It can enter ecosystems directly in industrial waste effluent (Kot-Wasik, 2005) or precipitate in particulate matter (Wuana and Okieimen, 2011). Unlike lead and mercury, the effects of cadmium poisoning are not well understood (Wuana and Okieimen, 2011). The heavy metal is highly persistent (Wuana and Okieimen, 2011) with a half-life of 10 to 30 years, posing delayed health risks when consumed in small quantities over a long period of time (Kot-Wasik, 2005). Between 1965 and 1985, Japanese medical researchers conducted studies on the effects of cadmium in the human body (Watanabe et al., 2004). *Itai-Itai*, or "ouch-ouch" disease was identified as the crippling bone and joint pain experienced by individuals with long-term exposure to cadmium (Tsuchiya, 1969).

1.2.3 Lead

Lead contamination stems from agricultural pesticide and fertilizer use (Wuana and Okieimen, 2011), industrial manufacturing of batteries, and transportation pollution (Kot-Wasik, 2005; Wuana and Okieimen, 2011). Historic use of lead-based paint may remain a source of lead contamination although regulated in both Canada (Health Canada, 2009; O'Grady and Perron, 2011) and the United States (CDC, 2014) since the 1970's. When inhaled or ingested, lead degrades neurological and gastrointestinal functions (Wuana and Okieimen, 2011). In the skeletal system, it can cause joint weakness (Wuana and Okieimen, 2011), calcium deficiencies (Dartmouth, 2012), and anemia (Kot-Wasik, 2005). Recent studies (Kot-Wasik, 2005) have found a decrease in cases of human lead poisoning in response to the phase-out of leaded gasoline. Since 1990, both the United States and Canada have banned the use of leaded gasoline under the Clean Air Act (EPA, 1996) and the Canadian Environmental Protection Act (Stephanson, 2009).

1.3 The Practicality of Spatial Interpolation

Sampling lake sediments to identify contamination is time consuming, and has become increasingly costly (Lui et al., 2006; Forsythe et al., 2010; De Solla et al., 2012; Simpson and Wu, 2014). Since surveying an entire lake basin for sediment samples is an unfeasible task, sediment surveys are designed to sample materials at specific locations throughout the study site. Contamination levels of unsampled sediments can be predicted using geostatistical methods of interpolation; a time and cost-effective solution to whole-lake sampling.

The fundamental theory underlying geospatial interpolation is Tobler's First Law of Geography, that is, "everything is related to everything else but near things are more related than distant things" (Tobler, 1970). In 1951, kriging spatial interpolation was developed by Danie G. Krige and later refined by mathematician Georges Matheron in 1973 (Largueche, 2006). Kriging is a geostatistical estimation of unknown values calculated from spatial weights between two or more measured points (Clark 1982; Ouyang et al., 2003; Hessel et al., 2007). The primary goal of kriging is to create a spatially continuous layer of information from isolated point data to better understand spatial phenomena (Forsythe et al., 2016b).

Predictions are made with a degree of uncertainty; herein lies the unexpected advantage of kriging over other methods of spatial interpolation such as Inverse Distance Weighting (IDW) (Willmott, 1984; Ziary et al., 2007). The accuracy of a predictive model is compared against the actual data set by systematically removing a single data point from the kriging analysis, running the predictive model, then comparing the predicted value to the actual data point removed (ArcGIS for Desktop, 2016). These errors are displayed as cross-validation statistics at the end of a kriging analysis, allowing the user to assess the validity of the interpolated surface created (Table 1-1). Cross-validation statistics are measured using the same units as the data source (in this case, $\mu\text{g/g}$). Conclusions drawn from interpolated surfaces with minimal prediction errors not only validate the scientific integrity of contamination surfaces produced through kriging, but also increase user confidence in knowledge production and decision making. For this reason, it is hard to question the inherent value of kriging spatial interpolation for its practicality, time and cost effectiveness.

The most valuable component of a kriging model is the semi-variogram (Figure 1-1; Krivoruchko, 2005; Forsythe et al., 2010). Semivariance is calculated as half of the squared difference between two points (Goovaerts, 1999; Lui et al., 2006). It is expected that correlation between two points will weaken as their distance apart increases (Hessel et al., 2007). This is identifiable by the range

of the semi-variogram from the sill, or where the maximum variance occurs in the data (McBratney and Webster, 1986). Semi-variograms are calculated for each contaminant, using several models. The spherical, Gaussian, and exponential models were assessed for their fit to the shape of the semi-variogram by Gawedzki and Forsythe (2012). Methods of assessing the model of best fit are traditionally biased (Hessl et al., 2007) when fitted by “eye” (Gotway and Hartford, 1996). A degree of expert knowledge of the data and study area is essential to select the most appropriate fit (McBratney and Webster, 1986).

Table 1-1: Cross-validation statistics, with conventional standards of accuracy (ArcGIS for Desktop, 2016).

Cross Validation Statistic	Conventional Standards	What is being measured?
Mean Prediction Error (MPE)	Close to 0	“The average difference between the measured and predicted values”
Average Standard Error (ASE)	Less than 20	“The average of the prediction standard errors”
Standardized-Root-Mean-Squared Prediction Error (SRMPE)	Close to 1	The degree of under- or overestimations in variability of predictions made
Minimal difference between the RMSPA and ASE	Minimal difference	How well the model predicts actual values vs. average prediction standard errors

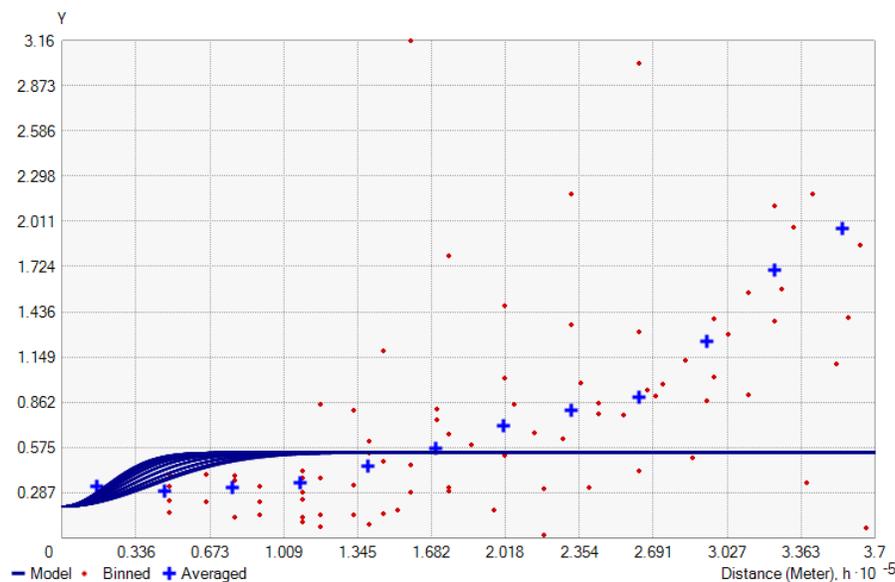


Figure 1-1: Example semivariogram displaying the fit of a Gaussian model through the 1971 Hg data points

1.4 Focus on Lake Erie

The Lake Erie watershed is the most densely populated and highly industrialized Great Lake (GLIN, 2004). The Regional Plan Association of the Tri-State area (RPA, 2008) has identified Lake Erie to be entirely encompassed by the emerging Great Lakes Megaregion (Figure 1-2). Based in part on the research of Jean Gottmann, *Megalopolis: The Urbanized Northeastern Seaboard of the United States* (1961), these megaregions have been developed by “the extraordinary dynamics [of] such an enormous and powerful concentration of people and activities” (Murphy, 1962).

Intensifying industrial and agricultural land use surrounding Lake Erie has prompted environmental concern since the mid-19th century (Evers et al., 2011). Modern anthropogenic activities including industrial and synthetic chemical production (Forsythe and Marvin, 2005; Kot-Wasik, 2005), metallurgy (Kot-Wasik, 2005; Habashi 2011), industrial waste and waste management (Wuana and Okieimen 2011; Dartmouth, 2012; Forsythe et al., 2016a), agriculture and agricultural fertilizing (Kot-Wasik, 2005; Wuana and Okieimen, 2011; Forsythe et al., 2016b), coal power production and fossil fuel combustion (Kot-Wasik, 2005; Wuana and Okieimen, 2011; Dartmouth, 2012) contribute to heavy metal pollution throughout the lake. Often confused as the most contaminated Great Lake (Adams and Stadler-Salt, 2016), Lake Erie contributes to the conditions of Lake Ontario, and beyond, through its interconnected waterways.

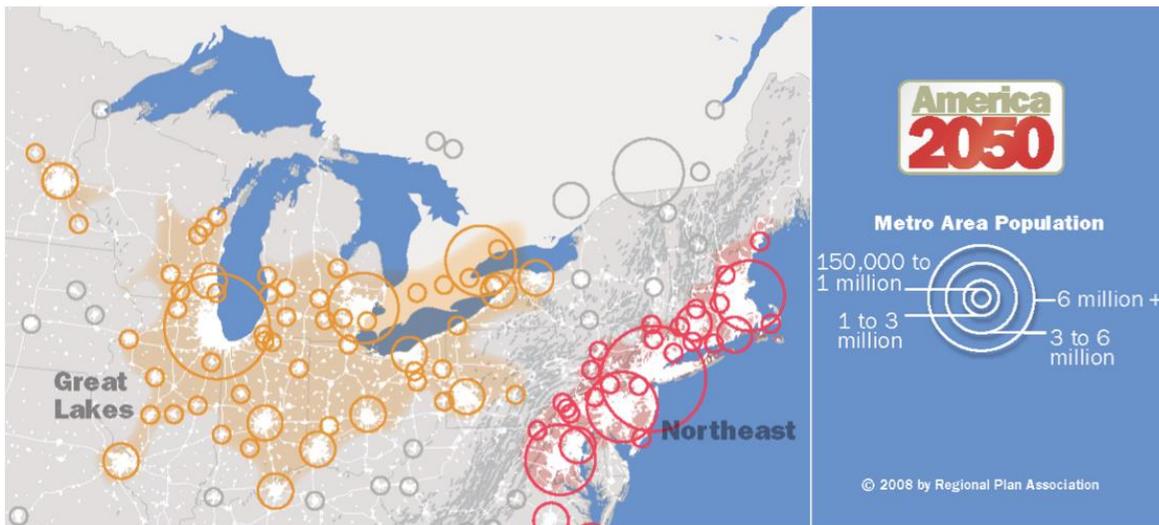


Figure 1-2: *The Emerging Great Lakes Megaregion encompasses the entire Lake Erie watershed, extending inland to several major American cities, and municipal regions of southwestern Ontario (RPA, 2008).*

Lake Erie was chosen as the study site for this dissertation for several reasons. First, recent sediment survey data were made available by the Canada Centre for Inland Waters (CCIW). With the addition of the 2014 dataset to existing historic surveys conducted in 1971, and 1997/1998, snapshots of changing contamination patterns can be produced over more than 40 years. Sampling densities and patterns between the survey periods decreased from 256 locations in 1971, to 55 locations in 1997/1998, and 34 points in 2014. Secondly, bathymetric data became available for the Lake Erie basin in 2014 from the National Oceanic and Atmospheric Administration (NOAA). The digital depth model allows for 3D visualizations of contamination patterns with additional geophysical context. Survey sample density, 3D geovisualizations, and multivariate analyses (including geophysical features and hydrodynamic processes) have yet to be analyzed for Lake Erie.

1.5 Research Objectives and Goals

This dissertation contributes to the growing body of knowledge on Great Lakes sediment contamination by producing new research findings in the fields of geostatistics, environmental contamination, and 3D geovisualization. The overarching theme of each individual study relates to enhancing scientific communication; not only methods of communication, but also the legitimacy of the science being communicated. To do so, this dissertation addresses the following research objectives and goals:

1. How many sampling points are required to produce statistically valid and spatially accurate predicted data surfaces using the kriging method of spatial interpolation?
2. How can contamination maps made from limited data be enhanced by auxiliary variables and alternative classification indices?
3. To develop complex yet approachable visual tools for pollution control efforts and overall knowledge production of sediment contamination.

Specific methodologies are derived for the empirical analysis of specific heavy metals within the context of Lake Erie. These research approaches may be adapted to additional organic or inorganic contaminants in other Great Lakes (where data are available). The purpose of this research does not extend to the explicit identification of contamination sources in Lake Erie. Nor does this study intend to propose remediation strategies to remove heavy metal contamination from the lake.

1.6 Chapter Summaries and Authorship Statements

This dissertation is composed of five chapters (Figure 1-3) containing three manuscripts of original research. The first chapter introduces both the research and study context for the entire dissertation. Manuscript style empirical research studies are presented in Chapters 2, 3, and 4.

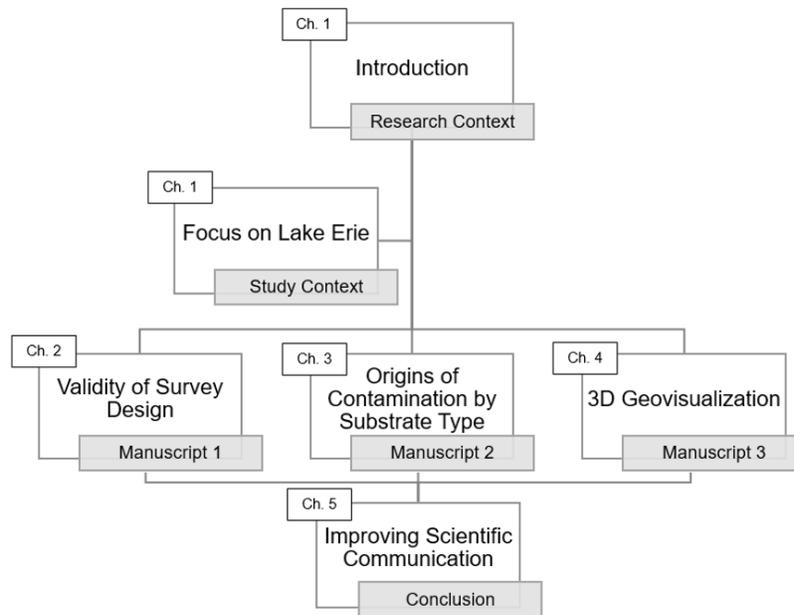


Figure 1-3: *Organizational structure of dissertation research.*

Each chapter employs kriging interpolation to address sediment contamination for one of three non-essential heavy metals (Chapter 2, mercury; Chapter 3, cadmium; Chapter 4, lead). Together, each chapter contributes to 1) the identification of lake-wide sediment contamination patterns; 2) improving the validity of statistically interpolated contamination surfaces; and 3) enhancing visualization using scientific communication tools. Chapter 5 concludes the dissertation by addressing future applications for each method of sediment contamination mapping throughout the Great Lakes basin.

1.6.1 Chapter 2: Assessing Statistical and Spatial Validity of Sediment Survey Design and Sampling Densities: Examples from Lake Erie

The minimum number of sampling points required to conduct a kriging interpolation analysis within Lake Erie is explored in this chapter. Based on changing survey point densities, the contamination patterns predicted by geostatistical interpolation of low sampling densities may not be as accurate as estimations made with high sampling densities; the question becomes, how low is too low?

The results of this study suggest a sampling density below which very generalized contamination patterns are predicted and from which pollution control efforts may not be justified.

Authorship Statement: Danielle Mitchell wrote Chapter 2 as a single author manuscript of original research under the supervision of Dr. K. Wayne Forsythe, and with editorial support from Dr. Chris H. Marvin and Debbie Burniston of Environment and Climate Change Canada (ECCC).

1.6.2 Chapter 3: Temporal Trends and Origins of Lake Erie Cadmium Contamination in Relation to Sediment Substrate Type Using Multivariate Kriging Analyses

This research study explores the benefits of multivariable kriging analyses in determining the origins of cadmium sediment contamination. Sediment contamination is identified to be low, moderate, or high as well as anthropogenically or naturally sourced. Auxiliary variables are shown to improve the predictive capabilities of spatially interpolated sediment contamination.

Authorship Statement: Danielle Mitchell wrote Chapter 3 as a single author manuscript of original research under the supervision of Dr. K. Wayne Forsythe, and with editorial support from Dr. Chris H. Marvin and Debbie Burniston of Environment and Climate Change Canada (ECCC).

1.6.3 Chapter 4: Multivariable 3D Geovisualization of Historic and Contemporary Lead Sediment Contamination in Lake Erie

Hydrodynamic processes and geophysical features are used in this study to enhance visualization tools for multiple-criteria environmental decision making. Advanced 3D geovisualization techniques are employed to create complex, yet accessible contamination maps to be used by a wide-range of stakeholders. The results of this study display barriers to sediment transport throughout the basins of Lake Erie, alluding to sediment transport fate over time.

Authorship Statement: Danielle Mitchell wrote Chapter 4 as a single author manuscript of original research under the supervision of Dr. K. Wayne Forsythe, and with editorial support from Dr. Chris H. Marvin and Debbie Burniston of Environment and Climate Change Canada (ECCC).

1.6.4 Chapter 5: Contributions to Great Lakes Research and Future Directions

The dissertation concludes by discussing its contributions to on-going Great Lakes research and subsequent scientific communication of research findings. Original research methodologies conceived in this thesis are adaptable to interpolating different sediment contaminants in other Great Lakes.

Chapter 2: Assessing Statistical and Spatial Validity of Sediment Survey Design and Sampling Densities: Examples from Lake Erie

Abstract: Spatial interpolation methods are used to translate randomized sediment contamination point data into informative area-based visualizations. Since 1968, the Canada Centre for Inland Waters has collected and analyzed sediment samples for organic and inorganic pollutants throughout the Great Lakes. In 1971, Lake Erie was first sampled using a square survey grid of 263 locations. Due to procedural costs, the most recent 2014 survey was reduced to 34 randomly selected sampling locations in deep offshore regions of the lake. Using the 1971 dataset, this study aimed to identify the minimum sampling density at which statistically valid, and spatially accurate predictions can be made using ordinary kriging. Through this process, it is possible to infer if the 2014 dataset includes enough points from which useful interpolated surfaces can be produced. Randomly down-sampled subsets at 10% intervals of the 1971 survey were created to include at least one set of data points with a smaller sample size than the 2014 dataset. Linear regression analyses of predicted contamination values assessed the spatial autocorrelation and goodness of fit between kriged surfaces created from the down-sampled subsets and the original dataset. Categorical and spatial analyses identified the numerical and spatial extent of over- and under-estimated contamination values. Down-sampled subsets at 10% and 20% of the original data density accurately predicted 51% and 75% (respectively) of the original dataset predictions. Subsets representing 70%, 80% and 90% of the original data density accurately predicted 88%, 90% and 97% of the original datasets predictions. Although all subsets were shown to be statistically valid, the spatial representation of contamination thresholds at sampling densities below 0.002 locations/km² are likely to create very generalized contamination maps from which policy and environmental decisions might not be justified.

Key words: Kriging, Great Lakes, Sediment Contamination, Geostatistics, Spatial Analysis, Visualization

2.1 Introduction

Harmful organic and inorganic pollutants from industrial, agricultural, and urban sources have contaminated the water, sediment, and ecosystems of Lake Erie since the mid-1800s (Evers et al., 2011). Non-essential heavy metals including mercury (Hg), cadmium (Cd), and Lead (Pb) are highly toxic and persistent throughout the ecosystems. Heavy metal poisoning can cause severe health risks to both human and wildlife populations through bioaccumulation, or direct contact with polluted water and sediments (Wuana and Okieimen, 2011).

The Lake Erie watershed is the most densely populated and highly industrialized Great Lake (GLIN, 2004), while also being the smallest by volume at 484 km³ (GLIN, 2016). Along its shoreline are several major cities including Toledo, Ohio, Cleveland, Ohio, Erie, Pennsylvania, and Buffalo, New York (Figure 2-1). Since the enactment of the Great Lakes Water Quality Agreement (GLWQA) in 1972, efforts have been made by both the Canadian and American governments, and environmental agencies to remediate historical and contemporary ecosystem pollution throughout the Great Lakes basin (IJC, 1972; IJC, 2002). Toxic substances are more persistent in the deepest parts of Lake Erie (in the eastern basin) which continue to contaminate the basin's ecosystems (Forsythe and Marvin, 2005; Forsythe et al., 2010). Historic and long-term monitoring and mapping of these contaminants is crucial for ongoing remediation efforts.



Figure 2-1: Lake Erie and its basins (Marvin et al., 2002; Esri, 2017).

2.1.1 Data Collection

Since 1968, the Water Science and Technology Directorate (previously the National Water Research Institute) and the Great Lakes Surface Water Surveillance Program have collected and analyzed sediment samples for organic and inorganic pollutants throughout the Great Lakes (Marvin et al., 2003; Marvin et al., 2004; Rukavina et al., 2013; Forsythe et al., 2016a). In 1971, sediment samples were collected from Lake Erie using a square survey grid. The survey grid was composed of 263 sampling points spaced approximately 10 km apart. The most recent Lake Erie sediment survey was conducted in 2014 where the sampling density was significantly reduced due to the increased cost of sampling procedures (Forsythe et al., 2010). Sampling points did not follow the original sampling grid used in 1971; instead a restricted number of sampling points were randomly selected in deep offshore regions of Lake Erie, placing emphasis on non-point source pollution. Only 34 sampling points make up the 2014 survey, almost eight times smaller than the initial dataset from 1971.

2.1.2 Geovisualization of Sediment Contamination

Dot maps (Figure 2-2) are a traditional medium used to display sediment contamination (Largueche, 2006; Forsythe and Marvin, 2009; Gawedzki and Forsythe, 2012). Contamination patterns can be observed in point form, however trends and underlying factors for contaminant distribution (i.e. bathymetry, and lake circulation patterns) are not recognizable or available to assist with interpretation (Forsythe et al., 2016c). Due to advances in geospatial computation, spatial interpolation methods such as kriging are now used to model estimated contamination in a continuous data layer (Goovaerts, 1999; Lark and Ferguson, 2004; Forsythe et al., 2016b; Forsythe et al., 2016c). Not only are interpolated surfaces cost- and time-effective (Simpson and Wu, 2014), they also work to translate disjunctive point data into communicable visualizations of spatial patterns (Hessl et al., 2007). Doing so also expands the communication of scientific research to a much broader audience. For example, geovisualizations of trends, patterns, and processes are significant pieces of information in policy development (Lark and Ferguson, 2004; Forsythe et al., 2016a; Forsythe et al., 2016b), pollution control and management (Wu et al., 2011; Khosh Eghbal, 2014), environmental science (Forsythe et al., 2016b; Forsythe et al., 2016c) and risk assessment (Markus and McBratney, 2001).

2.1.3 Kriging

Developed in 1951, kriging was initially used to predict ore reserves in mines (Largueche, 2006; Gawedzki and Forsythe, 2012). Since then, the interpolation method has been successfully

applied in other studies in the fields of soil studies (Gotway et al., 1996; Yan et al., 2007; Simpson and Wu, 2014), hydrology (Olea, 1984; De Solla et al., 2012), environmental management (Chang et al., 1998; Lui et al., 2006), zoology (Villard and Maurer, 1996; McKenney et al., 1998), forestry (Chen et al., 2012), horticulture (Dille et al., 2002) and those specific to contamination in the Great Lakes (Forsythe and Marvin, 2005; Forsythe et al., 2016c). There is however some disagreement among researchers as to the minimum survey samples required for kriging to produce statistically valid and meaningful interpolated surfaces (Chang et al., 1998; Li and Heap, 2008; Simpson and Wu, 2014).

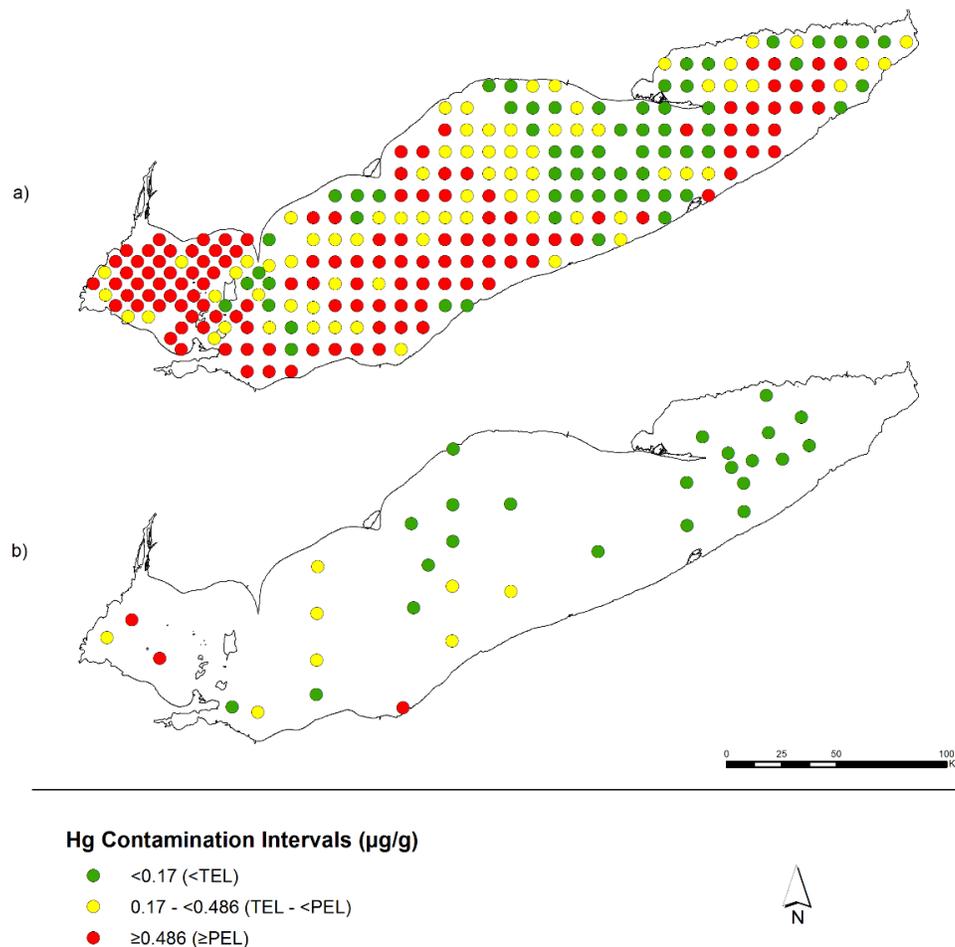


Figure 2-2: Categorical dot map of Hg sediment contamination from surveys of Lake Erie conducted in a) 1971 and b) 2014.

Regarding point count alone, Webster and Oliver (1993) recommend a minimum of 150, and up to 225 sampling points be required to conduct reliable spatial interpolations, whereas Burrough and McDonnell (1998) suggest between 50 and 100 sampling points. Kriging with less than 50

sampling points was determined to create “erratic” predictions (Webster and Oliver, 1993) and at 19, kriging predicts “results [which] may only be an artefact” of the actual process (Chang et al., 1998). All things considered, interpolation accuracies are not guaranteed to improve with an exhaustive sample density and remind us of the cost- and time-effectiveness of collecting limited survey samples (Simpson and Wu, 2014). Producing a statistically valid and meaningful interpolated surface is not only influenced by the number of sample points available (Webster and Oliver, 1993; Painter et al., 2001), but also the sampling density (Gotway et al., 1996; Ramsey, 1997; Chang et al., 1998, McKenney et al., 1998; Li and Heap, 2008), sampling pattern (Villard and Maurer, 1996; Goovaerts, 1999; Yan et al., 2007), and sampling direction (Clark, 1982; McBratney and Webster, 1986; Ouyang et al., 2003).

Sampling surveys in other fields of research have also been affected by the increased cost of conducting field work (Chang et al., 1998; McKenney et al., 1998; Lui et al., 2006; Yan et al., 2007; De Solla et al., 2012, Simpson and Wu, 2014). In the interest of prolonging current and future monitoring efforts, researchers have attempted to determine the smallest possible sampling densities (from which to conduct kriging interpolation) by down-sampling their original datasets into randomly, or systematically selected subsets (McKenney et al., 1998; Yan et al., 2007; Chang et al., 1998). Studies by Chang et al. (1998), McKenney et al. (1998), and Lui et al. (2006) were not all successful in producing reliable predictions in their respective experiments. Decreasing sampling density to 25% or less of the original data in McKenney et al. (1998) resulted in low prediction accuracies, whereas Chang et al. (1998) reported “sufficient” results from point densities reduced by 20% and 40% of the original dataset. There is not an absolute minimum number of samples from which to perform kriging. Expert knowledge of the study site and phenomena acting within its boundaries is also essential for determining the representational accuracies of an interpolated surface (Goovaerts, 1999; Forsythe et al., 2016b; Forsythe et al., 2016c).

Using the complete contamination dataset of Lake Erie from 1971, this study aims to identify the minimum sampling density at which statistically valid and spatially accurate predictions can be generated. The sampling pattern and density of the 2014 Lake Erie dataset is scrutinized to provide insight into whether it is adequate for performing spatial interpolation. This experiment intends to identify the threshold below which the sampling density of toxic contaminants in Lake Erie will produce very generalized prediction surfaces from which policy and environmental decisions might not be justified. In doing so, contemporary sediment surveys could be better

developed alongside historic survey designs to accommodate evolving research objectives over time to produce statistically valid and meaningful interpolated maps.

2.2 Data

Sediment sampling surveys of Lake Erie were conducted by Environment Canada in 1971, and the Great Lakes Surface Water Surveillance Program in 2014. Contemporary sampling locations were randomly selected within deep offshore regions of Lake Erie. These were identified as having the greatest threats for sediment contamination. When sampling locations were randomly selected close to historic sample sites, they were relocated to that location to allow for temporal comparisons between sites. Descriptive statistics for Hg from the 1971 and 2014 surveys are presented in Table 2-1.

Table 2-1: Descriptive statistics of Hg sediment contamination samples from Lake Erie.

Year	Sample Count	Minimum (µg/g)	Maximum (µg/g)	Average (µg/g)	Standard Deviation
1971	263	0.008	7.488	0.610	0.706
2014	34	0.004	0.777	0.159	0.164

Sediment samples were collected as mini-box cores; for a complete outline of this sampling procedure see Painter et al. (2001) and Marvin et al. (2002). Using this method, the top three cm of sediment was collected, stored in pre-washed jars and frozen for transportation to the lab for analysis (Painter et al., 2001). The top three cm of the lake bed is considered the active layer, where the greatest amount of chemical and biological activity occurs (Dalia et al., 2014). As the upper limits of the lake bottom, it is also a “substance exchange” zone between the lake sediment and the lake water (Dalia et al., 2014).

Sediment samples from 1971 measured Hg contamination between 0.008 µg/g and 7.488 µg/g. The maximum contamination recorded is an outlier of the dataset but was not removed since it would defy the principles of randomized sampling. Extreme sediment contamination values also identify highly polluted areas and should not be discarded or ignored when creating estimated contamination surfaces (Lui et al., 2006; Li and Heap, 2008; Wu et al., 2011).

2.3 Methods

Ordinary kriging is the most commonly used technique among kriging interpolators, namely for its valuable output of prediction errors and surfaces (Ouyang et al., 2003; Forsythe and Marvin, 2009). This method was used to estimate the amount of Hg sediment contamination at unsampled

locations throughout Lake Erie. The 1971 dataset was best modelled using a Gaussian distribution with a major range of 100,000 m, a minor range of 50,000 m and an anisotropic direction of 90°; the data were assumed to be anisotropic in nature due to the influence of lake currents and circulation patterns on contamination distribution (Forsythe and Marvin, 2005; Figure 2-3). These parameters were determined through experimentation using various models and evaluating the cross-validation statistics against the conventions for statistically significant distributions. Geographic weight was applied to a maximum of 5 and minimum of 1 neighbours. The same parameters were used on all down-sampled subsets to ensure consistency of the modelling technique.

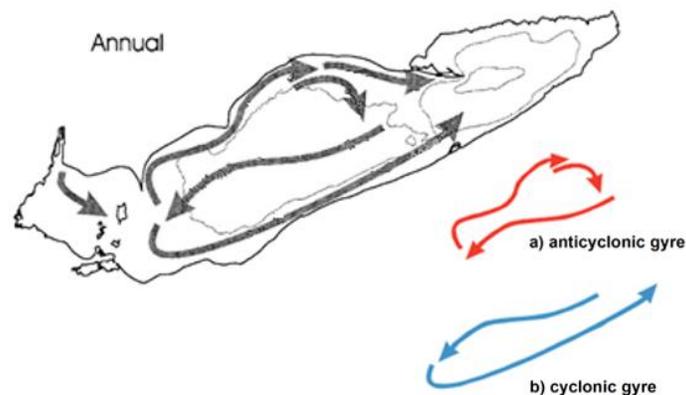


Figure 2-3: Annual Lake Erie circulation patterns adapted from Beletsky et al., 1999; the two-gyre system in the central basin is identified individually by the a) anticyclonic gyre and b) cyclonic gyre.

Cross-validation statistics assess the accuracy, or amount of error, between an estimated value and the actual value at a sampled location (Ouyang et al., 2003; Li and Heap, 2008). The resulting estimation errors indicate the strength of the model and semivariogram to predict all unsampled locations (McBratney and Webster, 1986; Forsythe and Marvin, 2005; Forsythe et al., 2010). The fit of a semivariogram model is accepted as accurate and unbiased when the Mean Prediction Error (MPE) is close to 0, the Average Standard Error (ASE) is as small as possible (<20), the Standardized Root-Mean-Squared Prediction Error (SRMSPE) is close to 1, and there is minimal difference between the MPE and ASE (Johnston et al., 2001; Simpson and Wu, 2014; Forsythe et al., 2016a; Forsythe et al., 2016b; Forsythe et al., 2016c). When the MPE is close to 0, there is a marginal average difference between measured and predicted values at the same location; a MPE greater than 0 indicates underestimation of the predicted values, whereas prediction errors less than zero show that values were estimated to be higher than they actually are (Osburn, 2000;

Forsythe and Marvin, 2005). Additionally, a SRMSPE greater than or less than 1 represent under- and overestimation of variability, respectively (Johnston et al., 2001; Forsythe et al., 2016b; Forsythe et al., 2016c).

2.3.1 Down-sampled Survey Density

The 1971 dataset is the most extensive survey of Lake Erie available for analysis, and for that reason acts as the ideal sampling density in this research (hereinafter referred to as the “complete dataset”). Since the 2014 dataset only includes 34 locations (13% of the 1971 survey) experimental subsets were down-sampled at 10% intervals to include at least one set of data points with a smaller sample size than the 2014 dataset. Down-sampled subsets (Figure 2-4) were selected by the stratified random sampling method using SPSS (IBM Corporation, 2013).

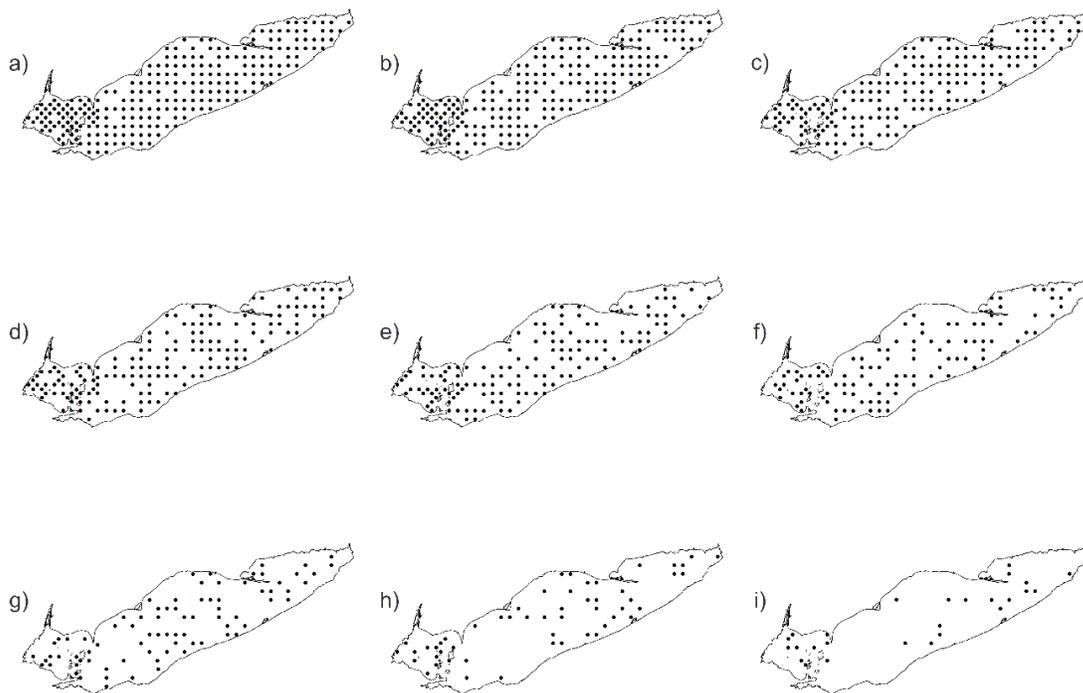


Figure 2-4: Point density maps of a) 90%, b) 80%, c) 70%, d) 60%, e) 50%, f) 40%, g) 30%, h) 20%, and i) 10% randomly down-sampled subsets from the complete dataset.

Random selection of sampling locations was intentional to replicate (as closely as possible) the 2014 survey design. Descriptive and spatial statistics of the complete dataset and 9 down-sampled subsets are presented in Table 2-2. The smallest down-sampled subset at 10% of the

complete dataset includes 26 points, 8 locations less than the 2014 dataset and is used as the lower sampling limit in this research.

Table 2-2: Descriptive statistics of Hg sediment contamination from the complete and down-sampled subsets.

Subset (%)	No. of Samples	Min. (µg/g)	Max. (µg/g)	Range (µg/g)	Average (µg/g)	Standard Deviation	Variance
100	263	0.008	7.488	7.480	0.610	0.706	0.498
90	237	0.008	7.488	7.480	0.620	0.722	0.521
80	210	0.010	2.929	2.919	0.594	0.579	0.335
70	184	0.008	2.929	2.921	0.559	0.558	0.311
60	158	0.008	7.488	7.480	0.616	0.775	0.601
50	132	0.008	2.929	2.921	0.632	0.564	0.318
40	105	0.008	7.488	7.480	0.714	0.879	0.773
30	79	0.008	2.929	2.921	0.552	0.563	0.317
20	53	0.008	1.881	1.873	0.520	0.476	0.227
10	26	0.056	2.817	2.761	0.873	0.779	0.607

Cross-validation statistics were analyzed for statistical validity of the kriging model and parameters used for each subset (Table 2-3). Predicted sampling values from each subset were compared against predicted sample values of the complete dataset using a linear regression analysis (Table 2-3). This process was performed in McKenney et al. (1998), where slopes and correlation coefficients of one identify identical predictions between two datasets. Slopes greater than one identify underestimation of the subset prediction to the complete dataset, and slopes less than one show that the down-sampled subsets over predict values in comparison to the complete dataset (McKenney et al., 1998).

2.3.2 Density Change Detection

Contamination maps utilizing the TEL and PEL were created from each down-sampled subset. The TEL refers to the concentration below which adverse biological effects are expected to occur rarely, while the PEL defines the level above which adverse effects are expected to occur frequently (CCME, 1999). The TEL and PEL for Hg are 0.17 µg/g and 0.486 µg/g, respectively. The maps were reclassified to represent the spatial extent of Hg values in three categories: <TEL, ≥TEL to <PEL, and ≥PEL (Table 2-4).

Each map was transformed into a raster dataset (500 m cell size, 0.05% of the 10 km grid) to perform calculations between the complete, and subset datasets. Contamination categories of the complete dataset were renamed as 1 (<TEL), 2 (≥TEL to <PEL), and 3 (≥PEL). Subset

categories were renamed as 100 (<TEL), 200 (≥TEL to < PEL), and 300 (≥PEL). The Raster Calculator tool in ArcGIS (Esri, 2017) was used to add the reclassified contamination categories of the complete dataset to each subset individually. This resulted in the creation of new numeric categories (Table 2-5) which identify the degree and location of change between the complete dataset and each down-sampled subset. Change categories 101, 202, and 303 represent no change in contamination classifications between the complete (original 1971) and subset datasets. Change categories 102, 103, and 203 identify classification discrepancies where subset datasets predicted lower Hg contamination values than was predicted by the complete dataset. In contrast, change categories 201, 301, and 302 identify classification discrepancies where subset datasets predicted higher Hg contamination values than was predicted by the complete dataset.

Table 2-3: Kriging cross-validation statistics of the complete dataset and all down-sampled subsets.

% Data	MPE (µg/g)	RMSPE (µg/g)	SRMSPE (µg/g)	ASE (µg/g)	RMSPE-ASE (µg/g)	Adjusted r ²	Slope
100	-0.001	0.580	1.216	0.479	0.101	n/a	n/a
90	0.001	0.611	1.255	0.489	0.122	0.996	0.988
80	0.001	0.384	0.909	0.423	-0.039	0.924	0.971
70	0.002	0.354	1.019	0.350	0.004	0.926	0.943
60	0.003	0.703	1.188	0.596	0.106	0.950	1.049
50	0.002	0.448	0.998	0.449	-0.001	0.873	1.046
40	0.000	0.907	3.138	0.292	0.616	0.726	0.602
30	-0.006	0.391	0.963	0.403	-0.011	0.799	0.866
20	0.005	0.367	1.058	0.364	0.002	0.847	1.202
10	-0.007	0.339	0.794	0.482	-0.105	0.908	0.754

Table 2-4: Hg contamination intervals, and respective thresholds, used for the density change detection analysis.

Contamination Intervals <TEL (µg/g)	Contamination Intervals ≥TEL to <PEL (µg/g)	Contamination Intervals ≥PEL (µg/g)
0.000 - <0.057	≥0.170 - <0.275	≥0.486 - <0.591
≥0.057 - <0.113	≥0.275 - <0.381	≥0.591 - <0.697
≥0.113 - <0.170	≥0.381 - <0.486	≥0.697

Table 2-5: Calculated contamination change categories and their meanings.

Change Category	Description
101	No change, <TEL
102	Subset (<TEL) underestimated complete dataset (\geq TEL to <PEL)
103	Subset (<TEL) underestimated complete dataset (\geq PEL)
201	Subset (\geq TEL to <PEL) overestimated complete dataset (<TEL)
202	No change, \geq TEL to <PEL
203	Subset (\geq TEL to <PEL) underestimated complete dataset (\geq PEL)
301	Subset (\geq PEL) overestimated complete dataset (<TEL)
302	Subset (\geq PEL) overestimated complete dataset (\geq TEL to <PEL)
303	No change, \geq PEL

2.3.3 Subset Kriging Analysis

All subsets had a MPE close to 0, SRMSPE close to 1, and low RMSPE and ASE. MPE values close to 0 showed a strong agreement between actual and predicted Hg contamination values, meaning the kriging model made reasonable estimations at unsampled locations. The highest SRMSPE values were present in subsets which randomly included the 7.488 $\mu\text{g/g}$ outlier (100%, 90%, 60%, and 40%). With SRMSPE greater than 1, down-sampled subsets at 90%, 60%, and 40% of the complete dataset had underestimated the variability of predictions made by the kriging model; not uncommon considering that kriging commonly underestimates predictions of high values (Webster and Oliver, 2007).

Estimated values from each subset were individually plotted against the estimated values from the complete dataset to further assess the correlation between prediction densities (slope and r^2 values presented in Table 2-3). Subset predictions were moderately (0.726) to strongly (0.996) correlated with predictions made using the complete dataset. At 60%, 50% and 20% of the complete dataset, slopes >1 identified underestimation of the subsets prediction to estimates made from the complete data. Despite the differences in data density, locations, and pattern, predicted values of each down-sampled subset were statistically valid according to cross-validation statistics and correlation coefficients.

Predicted Hg sediment contamination using the complete dataset is presented in Figure 2-5. Isolines demarcate the spatial boundaries of threshold effect levels (0.17 $\mu\text{g/g}$) and probable effect levels (0.486 $\mu\text{g/g}$) of Hg contamination (CCME, 1999).

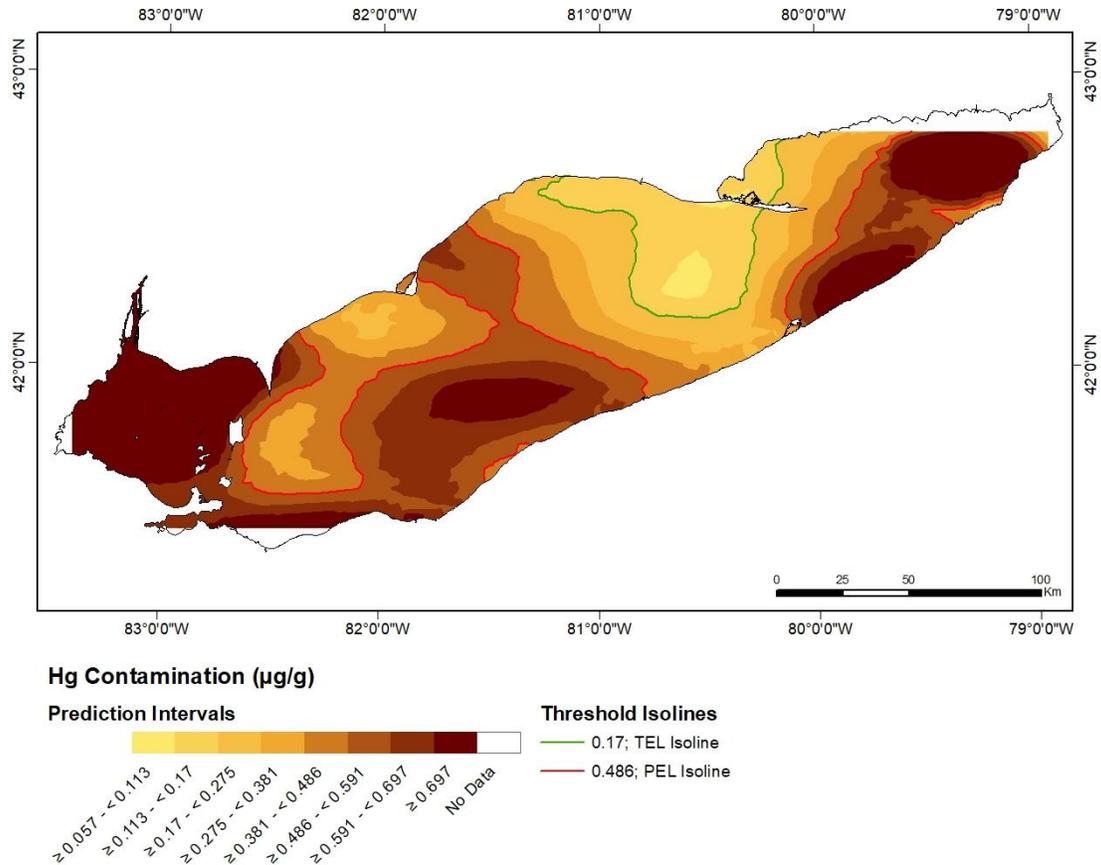


Figure 2-5: Kriging predictions of Hg sediment contamination in Lake Erie based on the complete dataset.

Most sediment contamination <TEL (2,679.72 km², or 10.39% of the Lake Erie analysis area) is located in the central basin, along the Ontario shoreline extending out into the centre of the lake. Sediment contamination \geq TEL and <PEL is located throughout the central and eastern basins representing 9,366.5 km², or 36.3% of the analysis area. Contamination \geq PEL represents the greatest spatial extent of sediment contamination (12,941.95 km², or 50.16% of the Lake Erie analysis area).

Subsets at 90%, 80% and 70% of the complete dataset produced very similar-looking Hg contamination patterns throughout Lake Erie (Figures 2-6a-c); exact spatial extents for all contamination categories are presented in Table 2-6. Due to the difference in spatial extent of varying point densities, Table 2-6 also lists the total analysis area for each subset (including the proportional representation of each subset to the full spatial extent of Lake Erie). Contamination <TEL remains located along the Ontario shoreline near the eastern extent of the central basin. Sediment contamination \geq TEL to <PEL is consistently found across the central basin into the

eastern basin; although the extent of this category grows towards the Niagara region at 80% and 70% subsets. The spatial extent of sediment contamination \geq PEL has consistent coverage in the western basin, as well as near the Ohio and New York shorelines. Overall, subset densities of 90%, 80% and 70% of the complete dataset represent highly comparable patterns of Hg contamination throughout Lake Erie.

Predictions made using 60% and 40% of the complete data density also produced contamination patterns similar to those predicted using the complete dataset as seen in Figures 2-7 (d,f) and Table 2-6. It is important to note, however, the unusually high SRMSPE of 3.138 (in comparison to the average SRMSPE of 1.254) produced using the 40% subset that highlights an underestimation of the 7.488 $\mu\text{g/g}$ outlier present in the predicted dataset. Isolines predicted using the 50% subset (Figure 2-7e) are especially abnormal for contamination $<$ TEL, as well as areas \geq TEL to $<$ PEL. These patterns are likely the result of a small sampling density which randomly omitted the 7.488 $\mu\text{g/g}$ outlier. Overall, the range of Hg contamination expands to include minimum contamination values of <0.057 $\mu\text{g/g}$ at these densities. The spatial extent of sediment contamination $<$ TEL closely resembles the same extent of $<$ TEL contamination as predicted by the complete data set, with the exception of the 50% subset. Conversely, contamination predictions throughout the western basin remain consistent with predictions made with subsets at greater sampling densities.

The overall spatial extent of prediction maps made from 30%, 20%, and 10% subsets are noticeably smaller (Figure 2-8g-i and Table 2-6). At 30% of the complete data density, the area of contamination \geq PEL is noticeably different from the complete dataset. Contamination \geq TEL to $<$ PEL represents most of the eastern basin, and encroaches on the expected \geq PEL monopoly in the western basin. At 20% of the complete PEL isolines demarcate increasingly more generalized contamination patterns in comparison to those created using the complete dataset. Due to the small sampling density in the 20% and 10% subsets, sediment samples were not located near the Ohio shoreline. This leaves the Sandusky and Cleveland areas of Lake Erie void of sediment contamination predictions. The 10% subset least resembles the contamination patterns created from kriging the complete dataset.

Table 2-6: Spatial extent of predicted Hg contamination by subset.

Subset	Hg Contamination	<TEL	≥TEL to <PEL	≥PEL	Total Analysis Area	No data
90%	Area (km ²)	2,497.62	9,489.27	13,001.47	24,988.36	811.63
	Area (%)	9.68	36.78	50.39	96.85	3.15
80%	Area (km ²)	2,423.21	10,285.14	12,186.13	24,894.48	905.52
	Area (%)	9.39	39.86	47.23	96.48	3.51
70%	Area (km ²)	2,881.99	10,295.58	11,810.79	24,988.36	811.63
	Area (%)	11.17	39.91	45.78	96.86	3.15
60%	Area (km ²)	2,137.80	10,983.36	11,772.25	24,893.41	906.59
	Area (%)	8.29	42.57	45.63	96.49	3.51
50%	Area (km ²)	760.12	10,883.36	13,250.99	24,894.47	905.52
	Area (%)	2.95	42.18	51.36	96.49	3.51
40%	Area (km ²)	3,292.16	8,833.13	12,727.37	24,852.66	947.34
	Area (%)	12.76	34.24	49.33	96.33	3.67
30%	Area (km ²)	2,413.04	11,426.09	9,447.82	23,286.95	2,513.06
	Area (%)	9.35	44.29	36.62	90.26	9.78
20%	Area (km ²)	3,255.45	12,071.81	8,815.36	24,142.62	1,657.38
	Area (%)	12.62	46.79	34.17	93.58	6.42
10%	Area (km ²)	2,106.66	5,628.08	12,434.61	20,169.35	5,630.65
	Area (%)	10.44	27.9	61.65	78.18	21.82

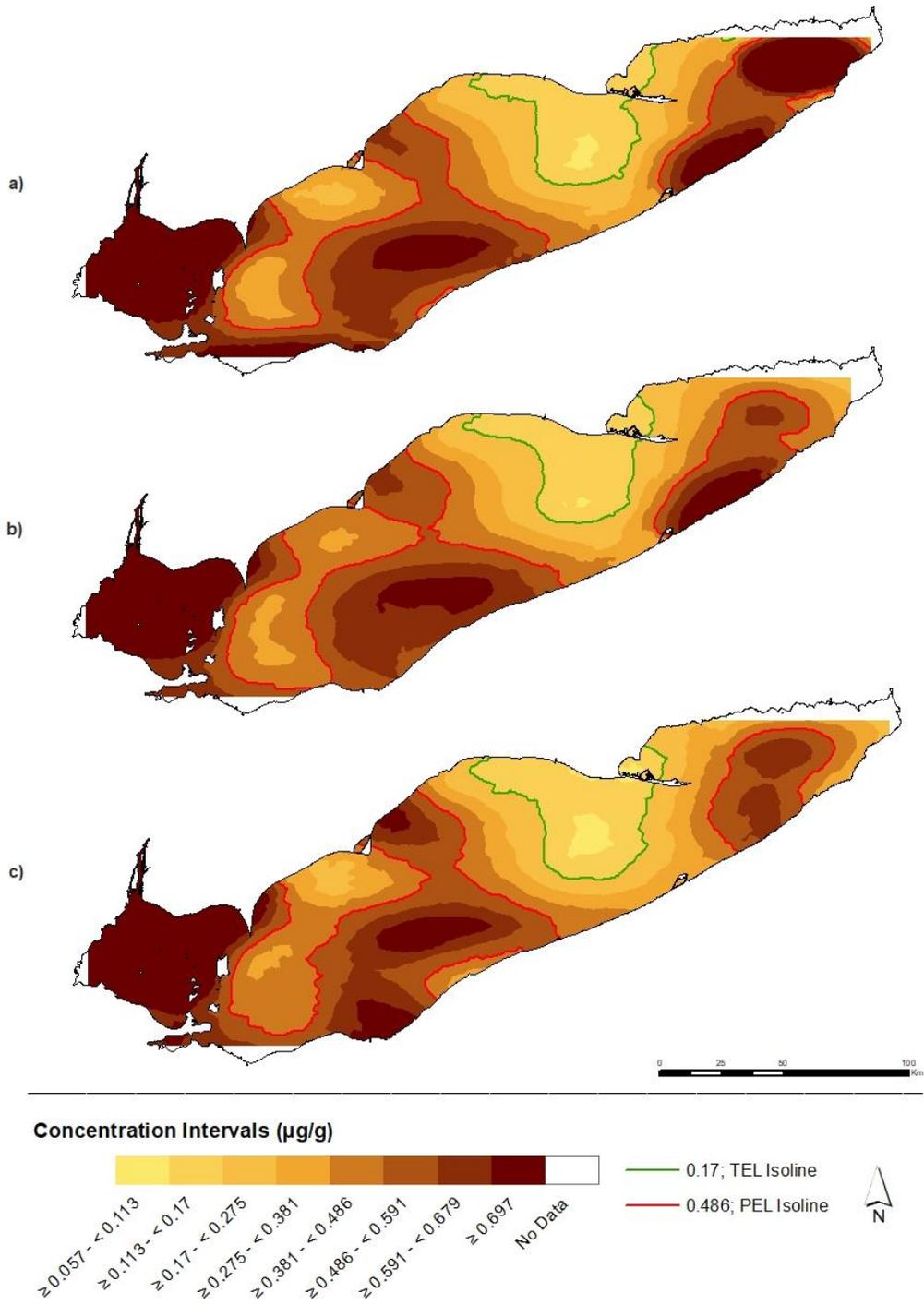


Figure 2-6: Kriging predictions of Hg sediment contamination in Lake Erie based on a) 90%, b) 80%, and c) 70% subsets of the complete dataset.

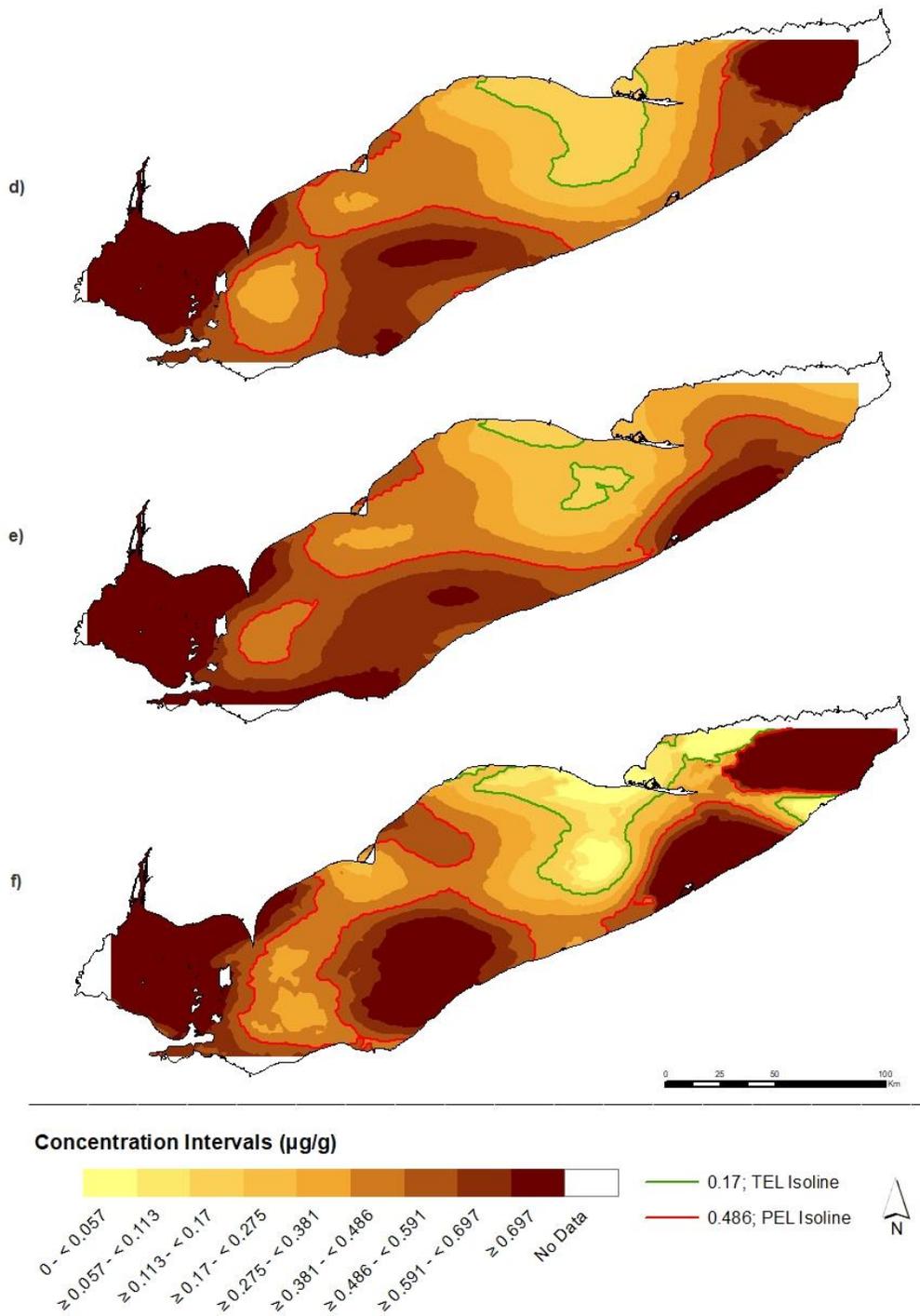


Figure 2-7: Kriging predictions of Hg sediment contamination in Lake Erie based on d) 60%, e) 50%, and f) 40% subsets of the complete dataset.

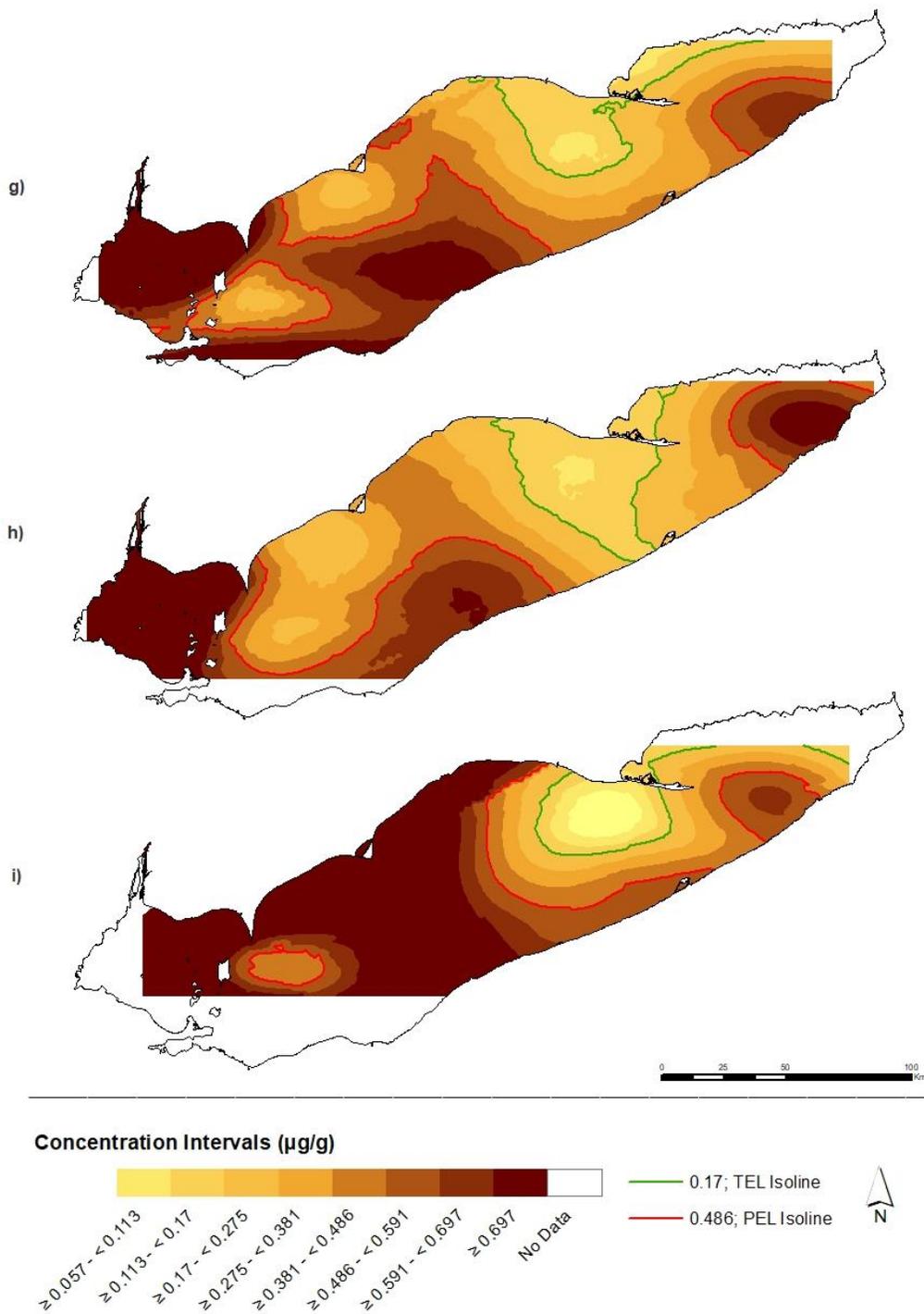


Figure 2-8: Kriging predictions of Hg sediment contamination in Lake Erie based on g) 30%, h) 20%, and i) 10% subsets of the complete dataset.

2.3.4 Categorical Change Detection Analysis

Kriging of Hg contamination was best predicted using the 90%, 80%, and 70% subsets. Less than 8% of pixels represented over- or underestimation of contamination thresholds. Change detection using the 90% subset reported no change in 97% of the pixels. Over- (2%), and underestimations (1%) were typically found at the boundary of two contamination thresholds; this pattern was common at the boundaries between $<TEL$, and $\geq TEL$ to $<PEL$ using 80%, and 70% subsets as well. Predictions made using the 80% and 70% subsets reported no change in 90% and 88% of pixels, respectively. Underestimations of $\geq TEL$ to $<PEL$, and $\geq PEL$ were common along the PEL isoline in the eastern basin, and nearest the Ohio shore. The locations and extent of over- and underestimations made by the 90%, 80%, and 70% (Figure 2-9a-c) subsets are likely representative of the transition zones between contamination thresholds and the size of the raster cell used to detect categorical change.

Greater discrepancies in predictions were made using the 60%, 50%, and 40% subsets (Figure 2-10d-f) than those at higher sampling densities; however, over- and underpredictions remained largely along the boundaries between contamination thresholds. Using the 60% subset, 86% of pixels reported seeing no change from predictions made using the complete dataset. Over- and underestimation of predictions were recorded in 6% and 8% of pixels respectively; most of which densely surrounded the contamination threshold boundaries. Predictions made using the 50% and 40% subsets predicted 75% and 79% agreement between predictions made using the complete dataset. The 50% subset overpredicted ($\geq TEL$ to $<PEL$) nearly the entire region identified as $<TEL$ by the complete dataset. Using this subset, over- and underestimations are no longer confined to contamination threshold boundaries as seen previously at higher sampling densities. Severe underestimation occurs using the 40% subset where a small section (147.5 km²) of Hg contamination is predicted to be $<TEL$ when previously identified as $\geq PEL$. An even smaller region (2.5 km²) of Lake Erie was classified to represent severe overestimation where the 40% subset predicted contamination $\geq PEL$ when previously reported as $<TEL$ in predictions made by the complete dataset.

Interpolation of the 30%, 20%, and 10% subsets (Figure 2-11g-i) resulted in the greatest instances of over- and underestimation of contamination versus the complete dataset. Predictions made using the 30% subset predicted 78% agreement between predictions made using the complete dataset. Agreements between datasets decreased to 75% using the 20% subset, and 51% using the 10% subset. Severe over- (269.75 km²) and underestimations (81 km²) were recorded in predictions made using the 10% subset. Over- and underestimations continued to be estimated

below the TEL isoline in the central basin, below the PEL isoline in the central basin above the PEL isoline in the eastern basin.

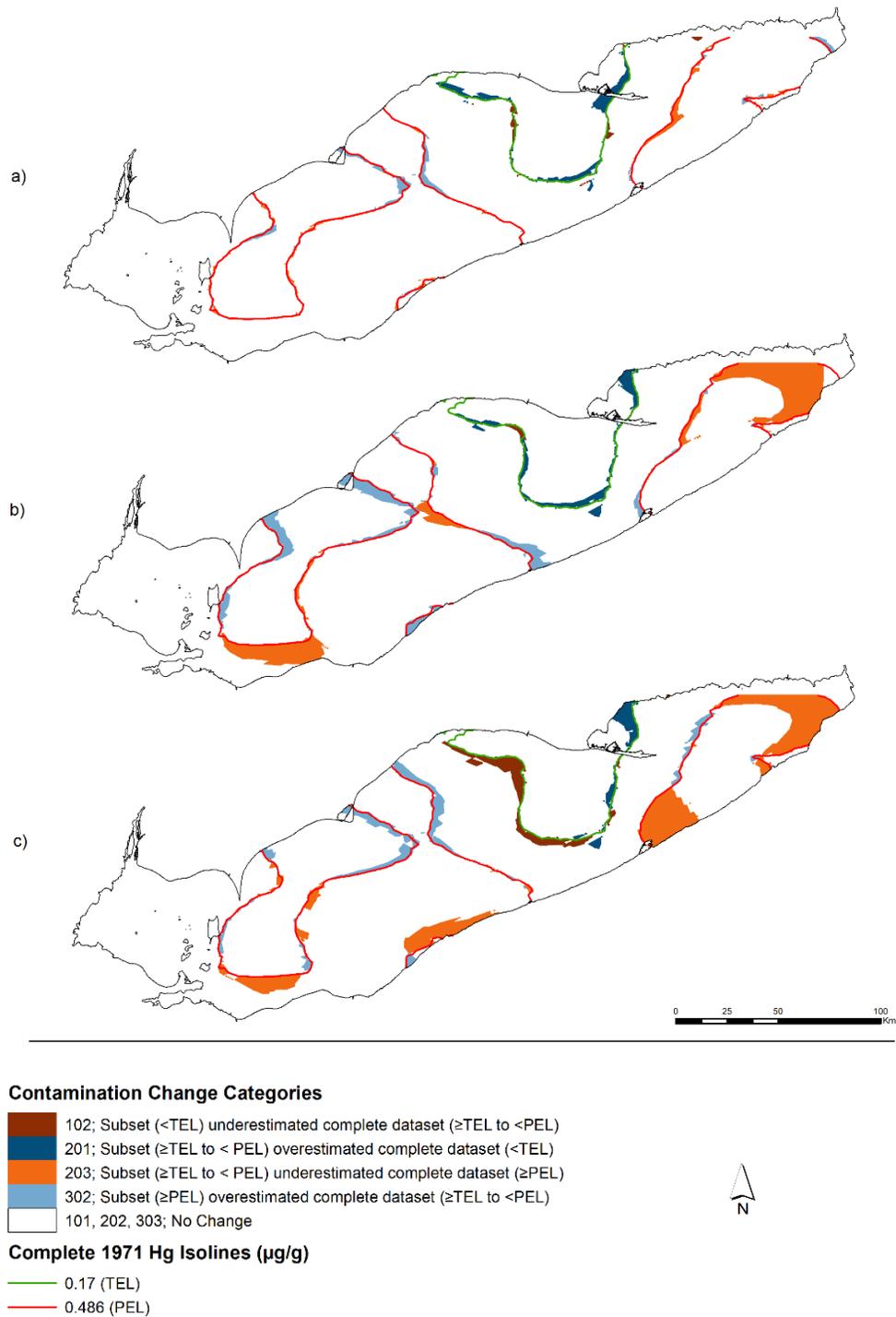
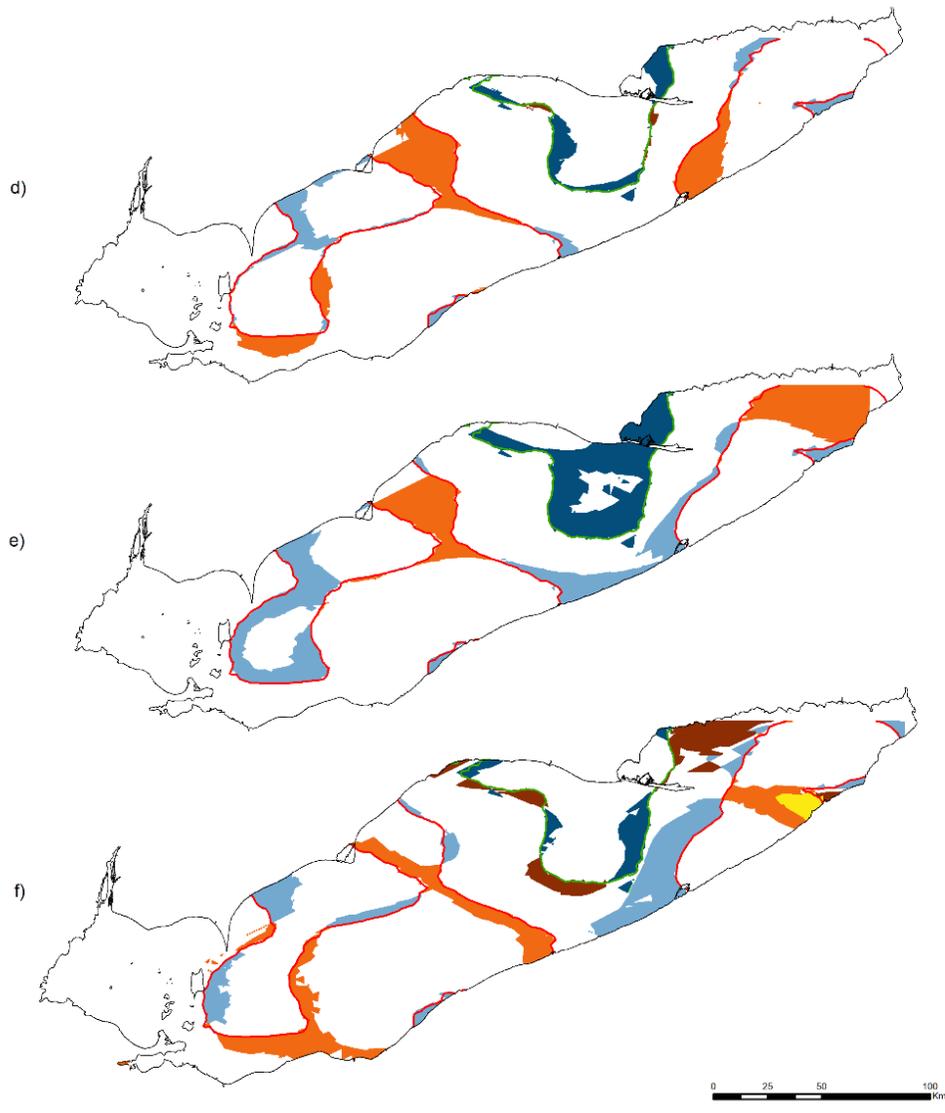


Figure 2-9: Change detection maps of no change, over-, and underestimations of Hg sediment contamination from kriging predictions based on a) 90%, b) 80%, and c) 70% subsets of the complete dataset.



Contamination Change Categories

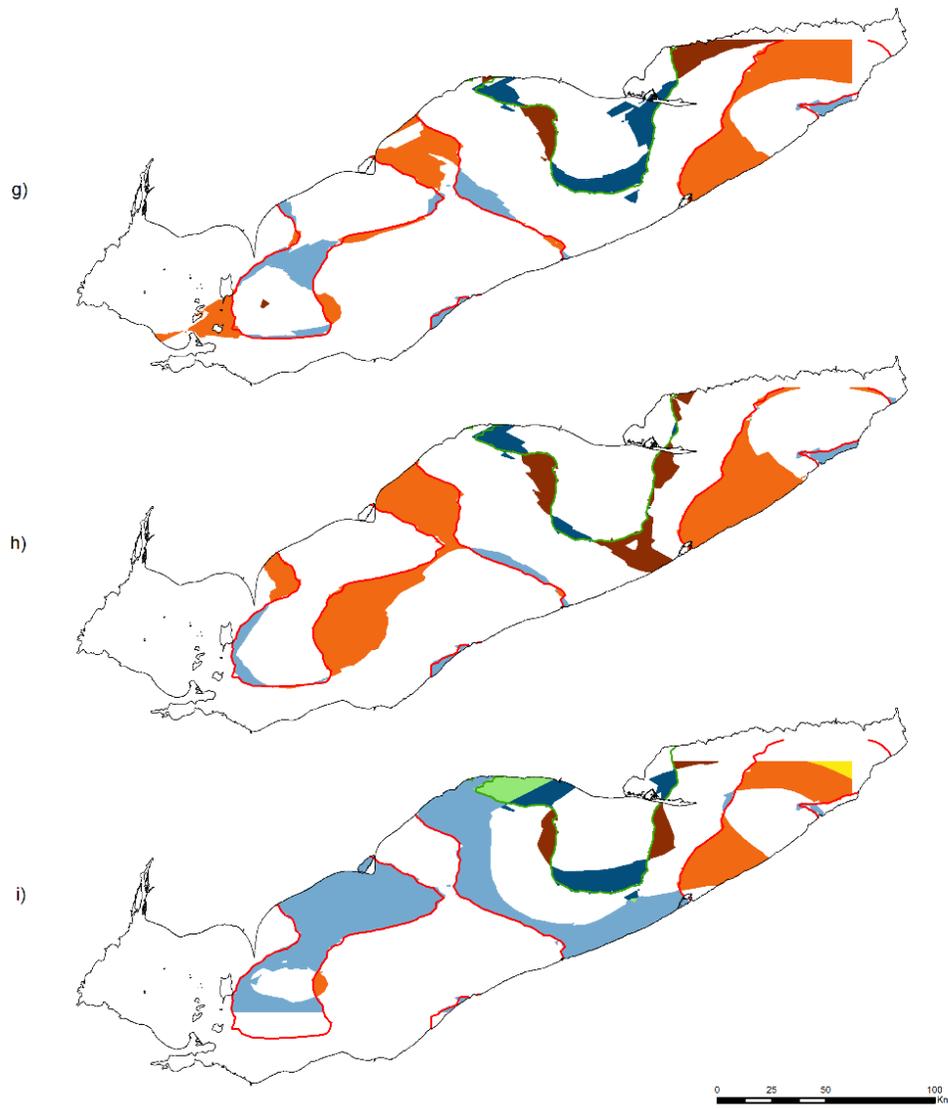
- 102; Subset (<TEL) underestimated complete dataset (\geq TEL to <PEL)
- 103; Subset (<TEL) underestimated complete dataset (\geq PEL)
- 201; Subset (\geq TEL to < PEL) overestimated complete dataset (<TEL)
- 203; Subset (\geq TEL to < PEL) underestimated complete dataset (\geq PEL)
- 301; Subset (\geq PEL) overestimated complete dataset (<TEL)
- 302; Subset (\geq PEL) overestimated complete dataset (\geq TEL to <PEL)
- 101, 202, 303; No Change

Complete 1971 Hg Isolines ($\mu\text{g/g}$)

- 0.17 (TEL)
- 0.486 (PEL)



Figure 2-10: Change detection maps of no change, over-, and underestimations of Hg sediment contamination from kriging predictions based on 60% (d), 50% (e), and 40% (f) subsets of the complete dataset.



Contamination Change Categories

- 102; Subset (<TEL) underestimated complete dataset (\geq TEL to <PEL)
- 103; Subset (<TEL) underestimated complete dataset (\geq PEL)
- 201; Subset (\geq TEL to < PEL) overestimated complete dataset (<TEL)
- 203; Subset (\geq TEL to < PEL) underestimated complete dataset (\geq PEL)
- 301; Subset (\geq PEL) overestimated complete dataset (<TEL)
- 302; Subset (\geq PEL) overestimated complete dataset (\geq TEL to <PEL)
- 101, 202, 303; No Change

Complete 1971 Hg Isolines ($\mu\text{g/g}$)

- 0.17 (TEL)
- 0.486 (PEL)



Figure 2-11: Change detection maps of no change, over-, and underestimations of Hg sediment contamination from kriging predictions based on 30% (g), 20% (h), and 10% (i) subsets of the complete dataset.

Smaller sampling densities and their spatial patterns can influence the coverage of predictions made throughout the lake. This change detection analysis also accounted for the proportion of

pixels representing “no data” as predicted surfaces made from subsets became increasingly smaller than that made from the complete dataset. The spatial extent of sampling points from the 90%, 80%, 70%, 60% and 50% subsets recorded “no data” in <1% of pixels. Prediction surfaces made using the 40%, 30%, 20% and 10% subsets recorded “no data” in 1%, 2%, 3%, and 19% of the study area, respectively.

2.4 Discussion

Prediction errors measure the legitimacy of interpolated surfaces. Although each down-sampled subset was deemed statistically valid by cross-validation statistics (Table 2-3), spatial analysis of each predicted surface identified significant deviations from contamination patterns of the complete dataset, which is considered to be the most accurate representation of Hg contamination in Lake Erie. In some cases, statistical validity may not translate into spatially reliable contamination maps. For example, while the cross-validation statistics of the 2014 Hg dataset are satisfactory (MPE of -0.009, RMSPE of 0.166, SRMSPE of 1.446, and ASE of 0.106), predictive contamination maps at sampling densities of 0.001 samples/km² may not produce reliable nor meaningful representations of contamination patterns throughout Lake Erie (Figure 2-12).

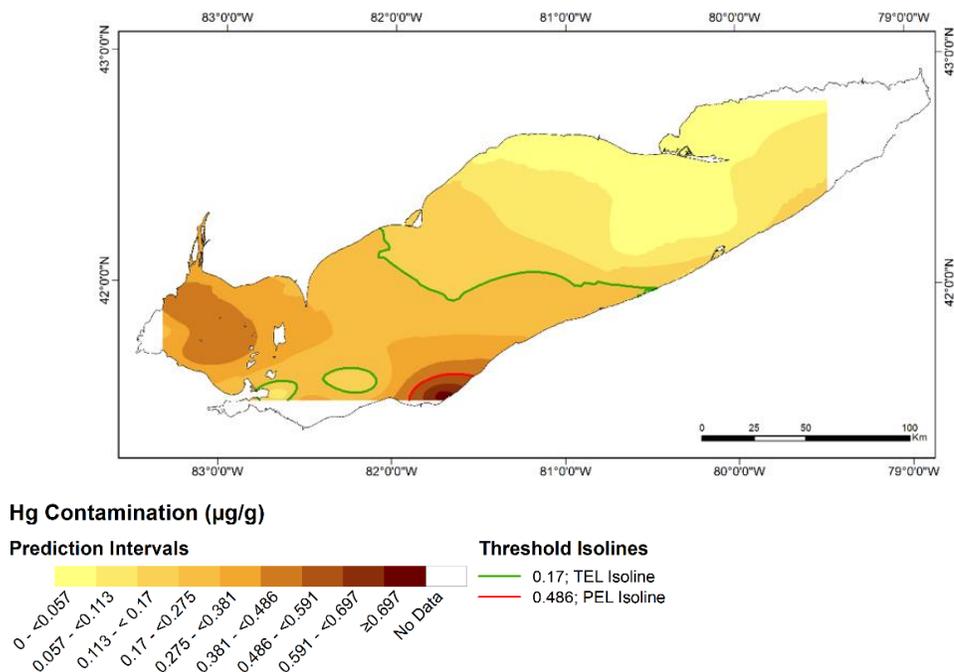


Figure 2-12: Kriging predictions of Hg sediment contamination in Lake Erie based on the 2014 dataset.

Herein lies some evidence regarding the potential confidence given by cross-validation statistics. Returning to the foundation of kriging, locations closer together are expected to record similar levels of contamination than those further apart (Verly et al., 1983). Sampling densities and location patterns are highly influential to the underlying data structure from which kriging is performed (Li and Heap, 2008). As the sampling density decreases in smaller subsets, the search neighbourhood expands much wider to find a maximum of 5 and minimum of 1 nearby sampling locations. The average distance between nearest neighbours in the 90% subset is <1km difference from the average nearest neighbours in the complete dataset. Nearest neighbours in the 10% subset are on average six kilometres further apart than those in the complete dataset. Sparse data density is known to produce more generalized interpolated surfaces (Li and Heap, 2008), which is noticeable in kriged maps created with subsets <20% of the complete dataset. Isolines demarcating the boundaries of sediment contamination <TEL, ≥TEL to <PEL and ≥PEL become smoother, altering their shape, size, and locations throughout the lake. Visualizations of Hg contamination in Lake Erie convey very different stories at subsets <20% of the complete dataset. From a decision-making perspective, the representations of sediment contamination, especially from the 10% subset could lead to them being of limited usefulness regarding the state of contamination in Lake Erie.

Exploratory data analysis revealed an outlier in the 1971 Hg contamination dataset (7.488 µg/g) which was not removed from the dataset when randomly down-sampling subsets were created. This is simply due to that fact that extreme outliers are particularly important pieces of information in contamination datasets (Li and Heap, 2008). The 7.488 µg/g outlier, located in the northeast corner of the eastern basin, was randomly selected to remain in the 90%, 60%, and 40% subsets. The relatively high sample density of the 90% subset did not allow the outlier to drastically influence the interpolation method nor the resultant contamination map. Cross-validation statistics for the 60% and 40% subset indicate slight underestimation of variability in their respective datasets (Johnston et al., 2001; Forsythe et al., 2016a) relative to the rest of the subsets. The categorical change detection analysis identified both severe underestimation (103; Subset (<TEL) underestimated complete dataset (≥PEL)) and severe overestimation (301; Subset (≥PEL) overestimated complete dataset (<TEL)) of predicted contamination using the 40% subset in the eastern basin. The outlier appeared to have no extraneous influence on predictions made using the 60% subset, producing similar contamination values as were made by the complete dataset. In subsets without the outlier, substantial portions of the eastern basin are underestimated (≥TEL to <PEL) in comparison to the complete dataset (≥PEL).

2.5 Conclusion

One goal of these experiments was to determine if the 34 samples collected in the 2014 survey are enough to produce statistically valid and spatially meaningful contamination maps. It is however not the number of samples but rather the density and spatial distribution of sampling locations which should have been questioned. The sampling density of the 2014 survey is approximately 0.001 samples/km², similar (albeit slightly higher) to that of the 10% subset of the 1971 dataset. Predicted contamination levels at this density likely created very generalized contamination maps in comparison to estimations made from the complete dataset.

This experiment with randomly down-sampled subsets showed how interpolated surfaces became increasingly generalized with smaller sampling densities. Increased procedural expenses have influenced survey design methodology for contamination sampling in Lake Erie. The 34-point survey from 2014 focused its resources on offshore regions of the lake in an effort to analyse non-point source pollution patterns. Low data densities created from biased data collection will likely produce exponential bias through data interpolation. Lake-wide spatiotemporal analyses were affected by the small sampling density of the most recent survey and it perhaps does not provide accurate interpolated sediment contamination patterns for the whole of Lake Erie from which environmental and policy decisions might be justified. Moving forward, a method to improve the survey design of sediment sampling could include smaller defined study areas within Lake Erie in which kriging could be performed at smaller sampling densities and surface areas. In addition, resampling at as many points as possible from the original 1971 survey grid may improve the reliability of contamination maps over time while applying greater emphasis on non-point source pollution analyses.

Chapter 3: Temporal Trends and Origins of Lake Erie Cadmium Contamination in Relation to Sediment Substrate Type Using Multivariate Kriging Analyses

Abstract: Using the kriging spatial interpolation method, Cadmium (Cd) sediment contamination in Lake Erie from 1971 (263 samples) and 1997/1998 (55 samples) were mapped as continuous data surfaces. To improve the interpolation accuracy between two different sampling densities, auxiliary variables including lake-wide sediment substrate types were integrated into the spatial analysis of Cd contamination patterns. The potential for adverse biological impacts posed by Cd sediment contamination to humans and aquatic life was measured by the threshold ($<0.6 \mu\text{g/g}$) and probable effect levels ($\geq 3.53 \mu\text{g/g}$) established by the Canadian Council of Ministers of the Environment. Patterns of Cd sediment contamination, and relationships with underlying sediment substrates were identified by overlaying threshold isolines from the 1971 and 1997/1998 interpolated surfaces. In 1971, 90.44% of Cd contamination above the probable effect level was located within regions dominated by fine grained sediments including mud (clay and silt), and sand. Contamination between the threshold, and probable effect level was also largely found in areas of the lake composed of mud (56.55%) while 20.85% of Cd in this range was found in glacial till (20.85%). By 1997/1998, contamination levels did not exceed the probable effect level; Cd contamination above the threshold effect level but below the probable effect level was again found predominately in mud substrates and glacial till. Contamination values were measured against naturally occurring background levels of Cd in Lake Erie ($<1 \mu\text{g/g}$) to distinguish anthropogenic sources of Cd pollution in the lake from naturally occurring traces. The index of geoaccumulation identified levels of contamination along the Long Point-Erie sill and the southwestern Ontario shoreline (composed predominately of glacial till) to be in the range of “practically unpolluted” by anthropogenic sources. The highest degree of anthropogenic pollution was found in regions of fine-grained materials (mud and silt). By combining Cd contamination surfaces with sediment substrate categories for analysis of anthropogenically derived pollution, regions of the lake at higher risk for long term sediment contamination can be targeted for pollution control and management action.

Key Words: Cadmium, Kriging, Great Lakes, Sediment Contamination, Geostatistics, Spatial Analysis

3.1 Introduction

Cadmium (Cd) is a naturally occurring heavy metal in the earth's crust at trace levels (Wester et al., 1992). Natural concentrations of Cd in soils range from 0.06 to 1.1 µg/g (USHHS, 2012), with a global average of 0.5 µg/g (Kabata-Pendias and Mukherjee, 2007). Particulate matter and debris resulting from forest fires, and volcanic eruptions contribute to natural inputs of Cd in the environment (Kabata-Pendias and Mukherjee, 2007).

Neither humans, flora, nor fauna require Cd for metabolic functions; one of the most toxic metals added into our environment from anthropogenic sources (Kabata-Pendias and Mukherjee, 2007). In the physical environment, Cd remains as a toxic heavy metal without biodegrading into "less harmful components" over time (Pachana et al., 2010). Long-term exposure or ingestion of Cd can result in serious human health complications (Kabata-Pendias and Mukherjee, 2007; Sabo et al., 2013), ecosystem degradation (Marvin et al., 2002; Pachana et al., 2010), as well as air, water, and soil pollution (Idriss and Ahmad, 2012; USHHS, 2012).

Anthropogenic activities contributing to Cd pollution include (but are not limited to): metallurgy (Wester et al., 1992; Kot-Wasik, 2005; Habashi 2011), industrial waste and waste management (Wuana and Okieimen 2011; Dartmouth, 2012; Forsythe et al., 2016a), mining and mine tailings (Kot-Wasik, 2005; Wuana and Okieimen 2011; Dartmouth, 2012), agriculture and agricultural fertilizer (Kot-Wasik, 2005; Wuana and Okieimen, 2011; Forsythe et al., 2016b), coal power plants, and fossil fuel combustion (Kot-Wasik, 2005; Wuana and Okieimen, 2011; Dartmouth, 2012). In 2000, the global use of Cd was 17,700 tons, 70% of which was used in the production of batteries (Kabata-Pendias and Mukherjee, 2007); a common industrial and household item with a poor history of proper disposal (USHHS, 2012).

3.1.1 *Fate of Cd Sediment Contamination in Aquatic Environments*

The total amount of Cd present in sediment samples (measured in µg/g) identifies the magnitude of heavy metal contamination at specific locations. This measure, although valuable in assessing sediment contamination, is not indicative of the mobility or bioavailability of Cd within its environment (Prokop et al., 2003; Idriss and Ahmad, 2012). Sorption and solubility characteristics of both heavy metals and sediment geochemistry determine the amount of Cd accessible for bioaccumulation throughout ecosystem components (Luoma, 1989; Adriaena et al., 2002; Prokop et al., 2003).

Influential geochemical variables include sediment grain size, pH, and the presence of clay and organic materials (Adriaena et al., 2002; Idriss and Ahmad, 2012). Bioaccumulation stemming

from sediment contamination is particularly concerning as concentrations typically exceed those in the water above by several orders of magnitude (Luoma, 1989; USHHS, 2012). Toxicity and mobility of Cd contamination in sediments best describe the heavy metals fate in aquatic environments (Pintilie et al., 2007; USHHS, 2012). As a water pollutant, Cd is known to be more mobile than other heavy metals (Idriss and Ahmad, 2012); as a sediment contaminant, the fate of Cd mobility is determined by geochemical conditions of the lake and lake bed (Pintilie et al., 2007).

3.1.2 Geochemical and Geomorphological Associations to Cd Contamination

Previous research has identified significant relationships between Cd and particular geochemical conditions of the contaminated sediments. Greater concentrations of heavy metal contamination (including Cd) are seen in fine-grained sediments (Lick et al., 1994; Adriaens et al., 2002; Marvin et al., 2002) near sources of industrial contamination (Marvin et al., 2002; Kabata-Pendias and Mukherjee, 2007). Naturally occurring Cd is often found in coarse-grained materials such as bedrock and glacial till (Zaidi et al., 2012). Across numerous studies of aquatic environments around the world, researchers identified significant Cd contamination in sediment grain sizes measured as clay (Thomas et al., 1976; Pintilie et al., 2007; USHHS, 2012; Birch et al., 2015). A relationship between Cd and pH of sediments was reported by Kabata-Pendias and Mukherjee (2007), the U.S. Department of Health and Human Services (USHHS, 2012) and Idriss and Ahmad (2012) although the nature of the relationship is contested amongst researchers. Kabata-Pendias and Mukherjee (2007) suggest high mobility of Cd in sediments at pH greater than five, whereas the USHHS (2012) reports the opposite, stating that low pH values in sediment increase the mobility of Cd contamination. No significant relationship is identified between the two variables by Idriss and Ahmad (2012).

3.2 Study Area and Data

Lake Erie sediments were first sampled for heavy metal contamination in 1971 by the Environment Canada Water Science and Technology Directorate (previously the National Water Research Institute) and the Great Lakes Surface Water Surveillance Program (Marvin et al., 2002; Rukavina et al., 2013). Due to the increasing resource requirements associated with collecting sediment samples, and more targeted sampling of specific offshore depositional areas, subsequent survey densities have decreased drastically since the initial assessment. The first sediment survey collected samples from 263 locations following a square grid pattern which decreased to 55 locations in the 1997/1998 survey.

In Lake Erie, the average natural concentration of Cd in lake sediments is <1 µg/g (Marvin et al., 2003). The mean Cd sediment contamination throughout the lake has exceeded this threshold since 1971 (Table 3-1).

Table 3-1: Descriptive statistics of Cd sediment contamination samples from Lake Erie.

Year	Min. (µg/g)	Max. (µg/g)	Mean (µg/g)	Standard Dev.	Variance
1971	0	10.8	2.36	1.48	2.2
1997/1998	0	4	1.23	0.93	0.87

3.2.1 Cd Contamination in Lake Erie

The Lake Erie watershed is entirely encompassed by the emerging Great Lakes Megaregion (RPA, 2008) developed by “the extraordinary dynamics [of] such an enormous and powerful concentration of people and activities” (Murphy, 1962). Intensive and haphazard urbanization, agricultural and industrial practices have been an environmental and ecosystem concern since the 1960’s (Ashworth, 1987). Detroit, Michigan and Cleveland, Ohio are known sources of historical industrial pollution to Lake Erie and its basin (Marvin et al., 2002).

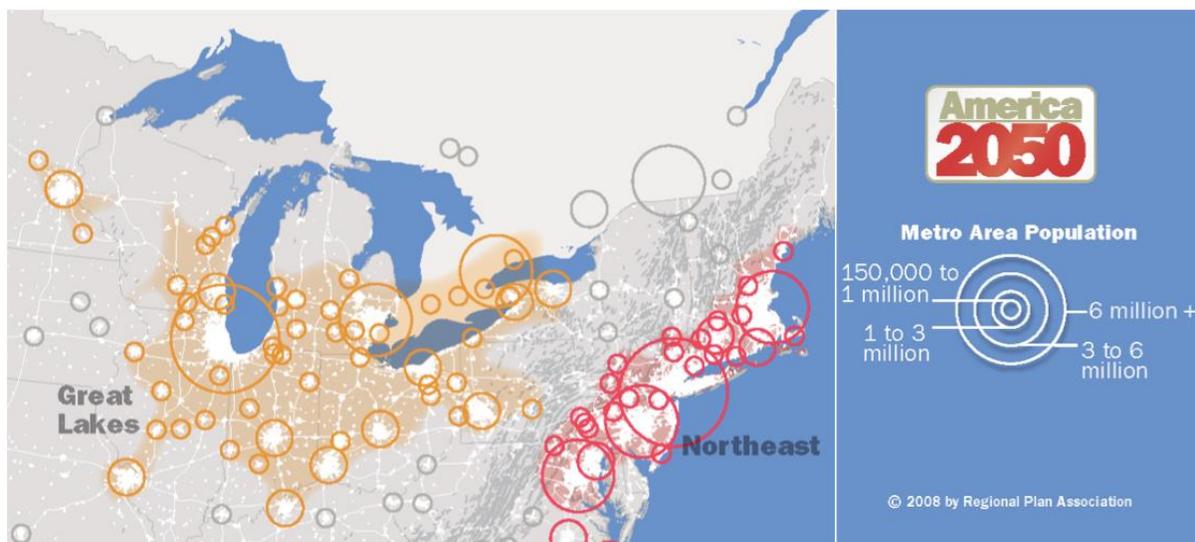


Figure 3-1: The Emerging Great Lakes Megaregion encompasses the entire Lake Erie watershed, extending inland to several major American cities, and municipal regions of southwestern Ontario (RPA, 2008).

3.2.2 Sediment Substrate Classification

The western, central and eastern basins of Lake Erie are unique glacial, and geomorphic features physically divided by glacial moraines (Thomas et al., 1976; Morang et al., 2011). Depths decrease offshore, and from the western to eastern basin, remaining relatively flat throughout the central basin (NGDC, 1999; Dusini et al., 2009). Thomas et al. (1976) describe a coarsening of

sediment grains from the western to eastern basins apart from fine-grained sediments from shoreline erosion located in the deepest regions of the eastern basin. This eastward trend in sediment coarsening can be attributed to the prevailing west/southwest winds across Lake Erie, and the amount of wind-wave energy created across the expansive central Basin (Thomas et al., 1976; Dusini et al., 2009). Lake Erie contains a wide range of sediment grain sizes including glacial till (0.004 to <256 mm), bedrock (≥ 256 mm), sand (62.5 μm to 2 mm), gravel (2 to <256 mm), clay and mud (0.98 to 3.9 μm) (Haltuch et al., 2000). Figure 3-2 displays a Lake Erie substrate map originally created by Haltuch et al. (2000) from data collected by Lewis (1966) and Rasul et al. (1997).

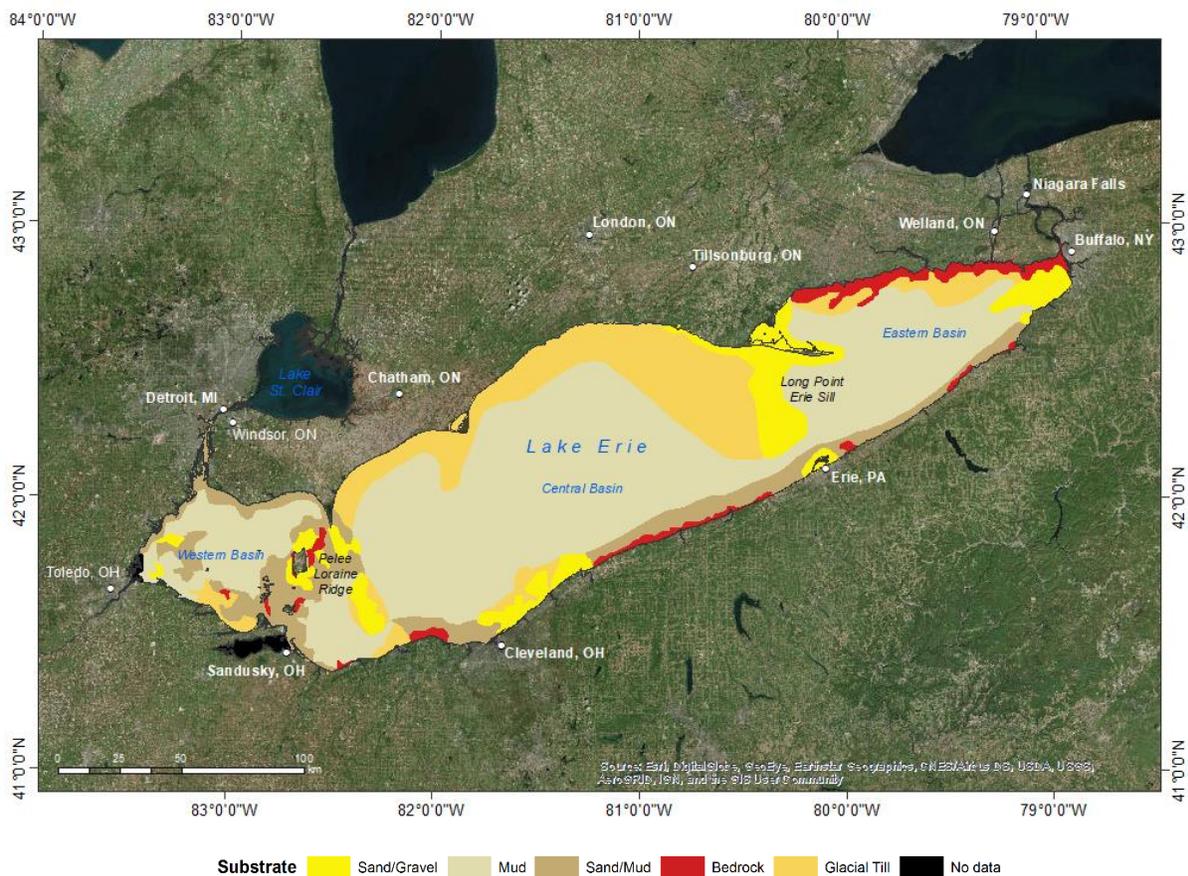


Figure 3-2: Lake Erie substrates, re-drawn from (Haltuch et al., 2000).

Lake Erie can be further categorized into depositional and transport zones. Depositional basins (outlined in Marvin et al., 2002) are similarly represented by substrate regions identified as “mud”. Fine sediment (0.98 to 3.9 μm) dominates the deep, offshore depositional basins (Thomas et al., 1976) including silts and clays, otherwise known as mud according to the Wentworth Scale of Grain Sizes (Table 3-2 - Haltuch et al., 2000). Increasingly larger grain sizes are in shallower

waters of the transport, or non-depositional zones (Thomas et al., 1976; Marvin et al., 2002) including the Pelee-Lorraine ridge, and the Long Point-Erie sill separating the western and central basins, and the central and eastern basins, respectively. Within seven kilometres of the shoreline is considered the littoral transport zone dominated by larger materials like sands, gravel, and glacial till (Thomas et al., 1976; Dusini et al., 2009; Morang et al., 2011). The distribution of substrates throughout Lake Erie is also indicative of lake-wide patterns of grain size and relative distribution. These spatial patterns are also representative of the relationships between grain sizes, lake circulation, wave-wind energy and overall movement of sediment types throughout the basin (Thomas et al., 1976; Morang et al., 2011).

Table 3-2: *The Wentworth (1922) Scale of Grain Sizes (Williams et al., 2011).*

Phi (ϕ)	Size (mm)	Wentworth Scale
≥ -8	256 mm	Boulders
-6 to -8	64 to <256 mm	Cobbles
-1 to -6	2 to <64 mm	Pebbles (granules to coarse)
4 to -1	62.5 μm to <2mm	Sand (very fine to coarse)
8 to 4	3.9 to <62.5 μm	Silt (very fine to coarse)
10 to 8	0.98 to <3.9 μm	Clay (mud)

3.3 Methods

This research aims to identify relationships between the presence of Cd in sediment samples and geochemical, and geomorphological characteristics of Lake Erie to enhance spatiotemporal models of lake-wide heavy metal contamination. Identifying the most vulnerable regions of Lake Erie to Cd contamination based on geochemical affinities of this heavy metal could contribute to a “weight of evidence” approach to lake-wide remediation and environmental policy implementation. Auxiliary variables to be included are the depth (m), pH, average grain size (ϕ), and proportional (%) substrate type of gravel, sand, silt, and clay. Regardless of the transport medium, Cd as a non-essential heavy metal contaminate in Lake Erie, must be tracked spatially and temporally to anticipate potentially harmful ecosystem conditions.

3.3.1 *Index of Geoaccumulation (I_{geo})*

Müller (1969) devised the first scale of measurement for heavy metal contamination in aquatic sediments (Gupta et al., 2014). The research was applied to contaminated sediments of the Rhine River (located in Western Europe) relative to water quality standards of the International

Association of Waterworks in the Rhine Catchment (IAWR); this index of geoaccumulation (I_{geo}) is now a widely accepted measure of sediment enrichment (Gupta et al., 2014; Birch et al., 2015).

From a single sediment sample, the index of geoaccumulation identifies the amount of anthropogenic contamination relative to the background, or natural presence of the same metal (Förstner et al., 1993). I_{geo} calculates sediment enrichment by:

$$I_{geo} = \log_2 (C_n / 1.5B_n) \quad [1]$$

where C_n is the measured concentration of metal contamination, B_n is the measured background concentration of the same metal, which is multiplied by a factor of 1.5 to correct for lithologic effects (Sabo et al., 2013; Gupta et al., 2014). Sediment cores, dating back before anthropogenic influences, are used to measure and identify the background concentration of specific metals (Birch et al., 2015). Background concentrations of Cd in Lake Erie were measured at <1 µg/g in all three basins (Marvin et al., 2003).

I_{geo} contamination values are ranked on a scale from 0 to 6, and then “graded” (Förstner et al., 1993) by the intensity of anthropogenic contamination (Gupta et al., 2014). The I_{geo} classification established by Müller (1981) is presented in Table 3-3. For reference, Förstner et al. (1993) determined the highest I_{geo} class (6) to be 100-fold greater than B_n .

Table 3-3: Index of geoaccumulation and pollution classifications (Müller, 1981).

Index of Geoaccumulation (I_{geo})	I_{geo} Class	Pollution Intensity
>5	6	<i>Very strong pollution</i>
>4-5	5	<i>Strong to very strong</i>
>3-4	4	<i>Strongly polluted</i>
>2-3	3	<i>Moderately to strongly</i>
>1-2	2	<i>Moderately polluted</i>
>0-1	1	<i>Unpolluted to moderate</i>
<0	0	<i>Practically unpolluted</i>

3.3.2 Canadian Sediment Quality Guidelines for the Protection of Aquatic Life (CCME)

The Canadian Council of Ministers of the Environment (CCME) have established thresholds of sediment contamination based on the likelihood of the concentration to create harmful biological conditions (CCME, 1999; Forsythe and Marvin, 2009). Measured Cd contamination <0.6 µg/g marks the Threshold Effect Level (TEL), at which harmful biological effects are not likely to occur

(CCME, 1999; Forsythe and Marvin, 2009). Measured Cd contamination $\geq 3.53 \mu\text{g/g}$ denotes the Probable Effect Level (PEL), where adverse biological effects are more likely to occur (CCME 1999; Forsythe and Marvin, 2009). Cd concentrations measured between the TEL and PEL should be of temporal concern, as the measurement may increase or decrease between sampling surveys.

3.3.3 *Co-Kriging*

An attempt was made to improve the interpolation accuracy of the 1997/1998 dataset using co-kriging, where only 55 samples were collected as opposed to 263 during the 1971 survey. Co-kriging is a variation of kriging where interpolation is conducted using auxiliary variables to enhance the accuracy of the estimated primary variable (Wackernagel, 1994; Zhang and Cai, 2015). Since this model includes a variety of independent variables, co-kriging is considered to be a multivariate problem (Genten and Kleiber, 2015). Genten and Kleiber (2015) warn that introducing numerous variables into a single spatial model can be challenging, especially when variables are disproportionately sampled (Wackernagel, 1994), and recorded at different units of measure. However, when strong autocorrelation is present between one or more independent variables to the dependant variable, co-kriging may be able to reduce the amount of variance within the interpolated model (Zhang and Cai, 2015). Conversely, Genten and Kleiber (2015) warn of insubstantial confidence in the degree of improvement using co-kriging over a single variable method of kriging.

Experiments were conducted with numerous combinations of independent variables (Cd and mean phi (ϕ), Cd and substrate types (%), Cd and depth (m), Cd and pH) however none of the auxiliary variables greatly improved the interpolation of Cd contamination in the 1997/1998 dataset as per Genten and Kleiber's (2015) suggestion. Since co-kriging yielded poor results, but the likelihood of correlation between Cd and geochemical variables provided was strongly suggested in previous research studies (Adriaena et al., 2002; Idriss and Ahmad, 2012), statistical analyses of relationships between Cd and auxiliary variables were conducted through alternative methods to co-kriging as suggested by Wackernagel (1994).

3.3.4 *Kriging*

The spatial interpolation tool, kriging, was used to create continuous data layers from Cd point data from the 1971, and 1997/1998 surveys. First, ordinary kriging was used to estimate the concentration of Cd in lake sediments at unsampled locations for both the 1971 and 1997/1998 datasets. The parameters used to model the distribution of the datasets for both years are

presented in Table 3-4. These were determined through experimentation, and goodness of fit based on estimation error outputs. The cross-validation statistics (Table 3-5) determine the amount of agreement between an interpolated value, and the measured value at the sample location (Ouyang et al., 2003; Li and Heap, 2008). Valid kriging models provide accurate interpolated values when the Mean Prediction Error (MPE) is near 0, the Average Standard Error (ASE) is smaller than 20, the Standardized Root-Mean-Squared Prediction Error (SRMSPE) is near 1, and the MPE and ASE are similar (Johnston et al., 2001; Simpson and Wu, 2014; Forsythe et al., 2016b). Predictive models with an MPE greater than and less than 0 represent under- and overestimation of kriged values in comparison to the actual measured values (Osburn, 2000; Forsythe and Marvin 2005). Underestimation of variability in the kriged surfaces are identified when the SRMSPE is >1, while a SRMSPE <1 represents an overestimation of variability (Johnston et al., 2001; Forsythe et al., 2016b).

Table 3-4: Kriging parameters used for modeling the 1971, and 1997/1998 datasets.

Year	Distribution Model	Major Range (m)	Minor Range (m)	Direction (°)	Lag Distance (m)
1971	Spherical	100,000	50,000	90	30,832
1997/1998	Gaussian	100,000	50,000	90	30,001

Table 3-5: Kriging cross-validation statistics of the 1971 and 1997/1998 sediment surveys.

Year	MPE (µg/g)	RMSPE (µg/g)	SRMSPE (µg/g)	ASE (µg/g)	RMSPE-ASE (µg/g)
1971	0.003	1.125	0.995	1.120	0.005
1997/1998	-0.009	0.933	0.996	0.938	-0.005

3.3.5 Change Detection

To quantify and assess the degree and spatial extent of change in Cd sediment contamination, a change detection analysis was employed between the 1971 and 1997/1998 interpolated surfaces. First, the 1971 interpolated layer was clipped to the extent of the 1997/1998 interpolation since the smaller sampling density of the 1997/1998 dataset resulted in different spatial interpolation extents. Next, to simplify the output change categories, Cd intervals for both time periods were aggregated to represent areas <TEL, ≥TEL to <PEL, and ≥PEL (Table 3-6). The contamination categories were then reclassified into unique numerical identifiers; for the 1971 dataset, <TEL was reclassified to 100, ≥TEL to <PEL to 200, and ≥PEL to 300. The 1997/1998 dataset was also reclassified as <TEL to 1, ≥TEL to <PEL to 2, and ≥PEL to 3. When added together using the

Raster Calculator in ArcGIS (Esri, 2017), unique combinations of the 8 numerical identifiers allowed for categorization of the output into degrees of Cd contamination change between 1971 and 1997/1998. All possible change combinations are listed in Table 3-7.

Table 3-6: Cd values organized by whole contamination intervals.

Contamination Intervals <TEL (µg/g)	Contamination Intervals ≥TEL to <PEL (µg/g)	Contamination Intervals ≥PEL (µg/g)
0 - <0.20	≥0.6 - <1.58	≥3.53 - <4.51
≥0.2 - <0.40	≥1.58 - <2.55	≥4.51 - <5.48
≥0.4 - <0.60	≥2.55 - <3.53	≥5.48

Table 3-7: Calculated contamination change categories and their meanings.

Change Category	Description
101	No change, <TEL
102	<TEL in 1971 to ≥TEL to <PEL in 1997/1998
103	<TEL in 1971 to ≥PEL in 1997/1998
201	≥TEL to <PEL in 1971 to <TEL in 1997/1998
202	No change, ≥TEL to <PEL
203	≥TEL to <PEL in 1971 to ≥PEL in 1997/1998
301	≥PEL in 1971 to <TEL in 1997/1998
302	≥PEL in 1971 to ≥TEL to <PEL in 1997/1998
303	No Change, ≥PEL

3.3.6 Statistical analysis

A multivariate regression analysis was calculated to assess the relationship between Cd sediment contamination and indicator variables at each sampling location (a method previously used by Idriss and Ahmad, 2012 and Sabo et al., 2013). In this case, correlations ($p < 0.05$) between Cd and indicator variables may allude to the fate of Cd sediment contamination relative to geochemical characteristics of Lake Erie.

3.3.6.1 1971 Regression

Regression of the 1971 dataset produced an adjusted r^2 value of 0.389 where $p < 0.001$. In this analysis, Cd was significantly ($p < 0.05$) correlated with pH, sand (%), silt (%), clay (%), and mean grain size (ϕ). The Durbin-Watson value (Montgomery et al., 2001) of 1.085 is indicative of positive autocorrelation of Cd contamination patterns throughout the lake. Between auxiliary variables, issues of moderate (>4) and severe (>10) multicollinearity (PSU, 2017) appear between the silt (%), clay (%) and mean grain size (ϕ) variables (variance inflation factor (VIF) of 4.803, 13.66, and 17.77, respectively). This redundancy was expected since clay and silt comprise a

large expanse of substrate type in Lake Erie, and are measured from 4 to 10 on the ϕ scale; this overlap prompted the removal of the mean phi variable from the regression analysis.

The regression analysis was re-run with all indicator variables previously used, except for mean grain size (ϕ) resulting in a marginally different adjusted r^2 of 0.377 where $p < 0.001$, and zero instances of multicollinearity ($VIF < 2$). Positive autocorrelation was identified by a Durbin-Watson value of 1.069, however the amount of explained variation was quite low. In this analysis, Cd was significantly influenced by clay ($\beta=0.504$, $p < 0.001$), silt ($\beta=0.162$, $p < 0.05$), and depth ($\beta=-0.244$, $p < 0.001$), where depth was the only variable indicating a significant negative relationship with Cd. Pearson correlations are presented in Table 3-8.

Table 3-8: Correlations between Cd values from 1971 and auxiliary variables.

Pearson Correlation	Depth (m)	pH	Gravel (%)	Sand (%)	Silt (%)	Clay (%)
Cd	-0.072	-0.310*	-0.115	-0.394*	0.329*	0.554*

*significance < 0.05

3.3.6.2 1997/1998 Regression

Regression of the 1997/1998 dataset including the mean phi variable produced an adjusted r^2 value of 0.198 where $p < 0.05$. Extreme multicollinearity ($VIF > 10$) was identified in the mean grain size (ϕ) ($VIF=150.924$), clay ($VIF=137.805$), silt ($VIF=17.203$), and sand ($VIF=8.113$) variables due to their redundancies in correlating mean grain size to sediment type. Again, the regression analysis was re-run omitting the mean grain size (ϕ) variable reducing all variables VIF to < 4 . Removing the redundant variable also changed the adjusted r^2 to 0.202 where $p < 0.05$ while maintaining positive autocorrelation throughout the dataset (Durbin-Watson of 0.563). There are no significantly influential variables to report (based on a significance threshold of $p < 0.05$). Variables reporting potential significance included clay ($\beta=0.248$, $p= 0.177$), silt ($\beta=-0.303$, $p= 0.055$), and depth ($\beta=-0.354$, $p= 0.014$) just as in the 1971 regression analysis. Pearson correlations are presented in Table 3-9.

Table 3-9: Correlations between Cd values from 1997/1998 and auxiliary variables.

Pearson Correlation	Depth (m)	pH	Gravel (%)	Sand (%)	Silt (%)	Clay (%)
Cd	-0.256*	-0.011	-0.361*	-0.261*	-0.106	0.188

*significance < 0.05

3.4 Spatial Analysis

Three areas of elevated contamination were identified from interpolation of the 1971 Cd survey (Figure 3-3). Along the southern extent of the central basin, two areas off the shores of Cleveland and Ashtabula, Ohio were identified to contain Cd sediment contamination $\geq 3.53 \mu\text{g/g}$ or $\geq \text{PEL}$. Even greater concentrations of Cd were predicted from the lower reaches, and mouth of the Detroit River discharging into the western basin of Lake Erie. Cd contamination in this location is likely to exceed $5.48 \mu\text{g/g}$, more than 1.5x the PEL. In total, Cd contamination $\geq \text{PEL}$ is estimated to represent $3,246 \text{ km}^2$, or 13% of Lake Erie sediments within the study boundaries. Moderate levels of sediment contamination were estimated in over 80% of the study area, throughout $21,277.75 \text{ km}^2$ of Lake Erie. Cd concentrations from 0.6 to $3.53 \mu\text{g/g}$ dominate most of the central basin, and the entirety of the eastern basin. Along the portions of the Pelee-Loraine ridge, and the Long Point-Erie sill are the lowest estimated concentrations of Cd contamination ranging from 0.6 to $< 2.55 \mu\text{g/g}$.

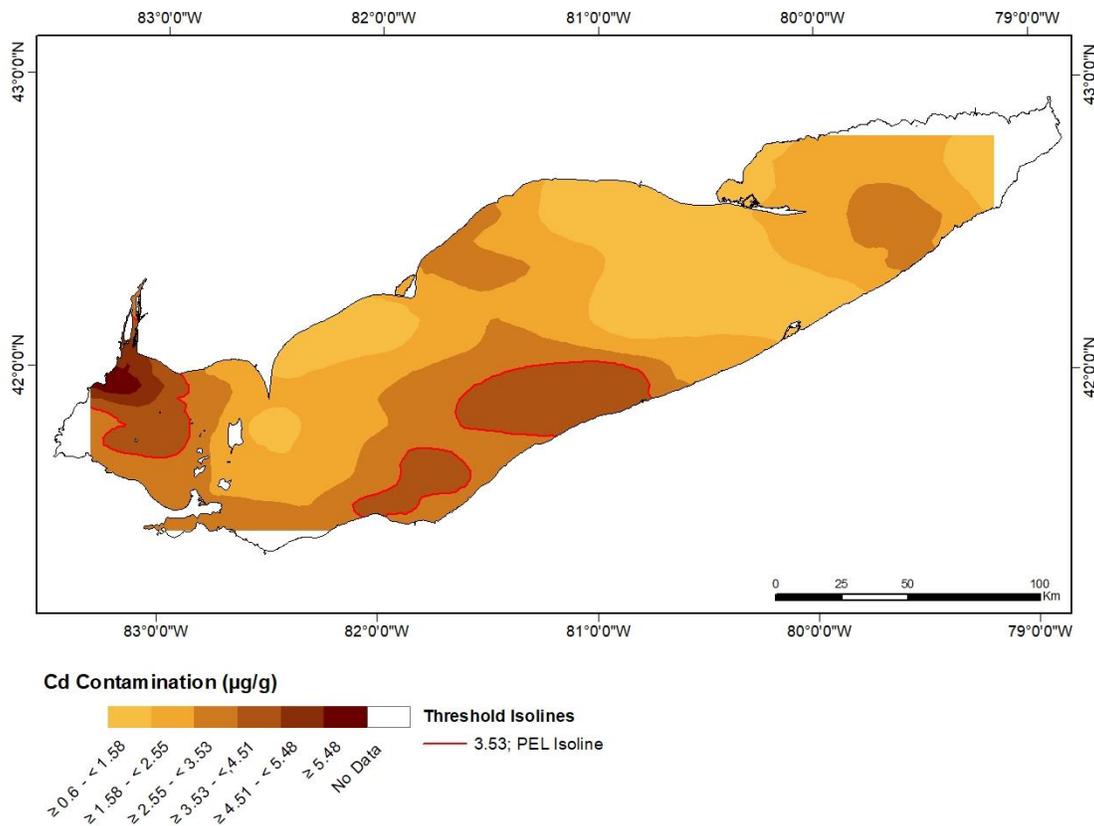


Figure 3-3: Kriging results of the 1971 Cd contamination in Lake Erie.

The 1997/1998 sediment survey predicted overall lower Cd contamination values across the lake (Figure 3-4). Nearly 90% of the lake's study area was estimated to have contamination values \geq TEL to $<$ PEL; the range of Cd values \geq TEL to $<$ PEL in the 1997/1998 contamination map do not exceed 2.55 $\mu\text{g/g}$ in contrast to the maximum interval measurement of $<3.53 \mu\text{g/g}$. Sediment contamination $<$ TEL was identified in 2,581 km^2 , or 10.5% of the study area, exclusively in the eastern basin. Interestingly, the southern shore of the central basin experienced higher concentrations of Cd nearest Cleveland and Ashtabula, Ohio, than any other area of the lake. This spatial pattern was also visible in the distribution of sediment contamination from the 1971 survey, although clearly not of the same magnitude in 1997/1998.

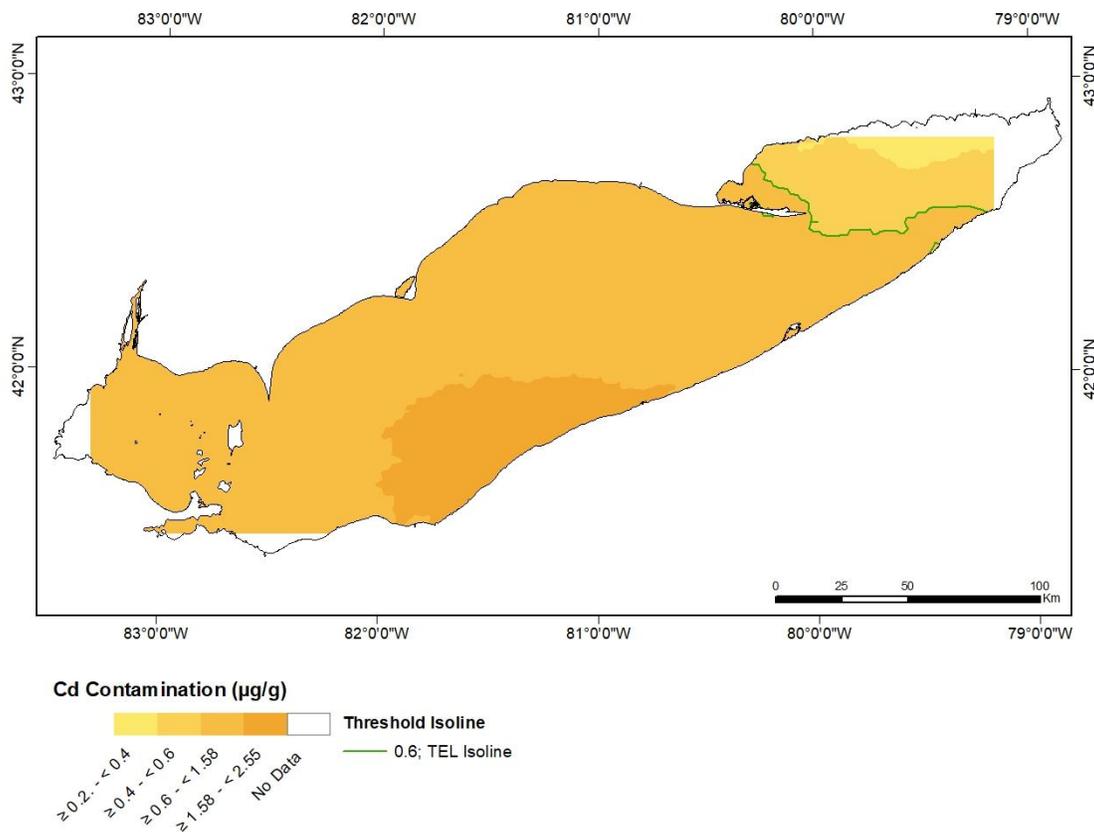


Figure 3-4: Kriging results of the 1997/1998 Cd contamination in Lake Erie.

The result of detecting spatiotemporal change in sediment contamination between 1971 and 1997/1998 is presented in Figure 3-5. The seemingly simplistic result was to be expected considering the contamination patterns presented in the 1971 and 1997/1998 maps. Areas of elevated sediment contamination identified from the 1971 survey were reduced to Cd values measuring \geq TEL to $<$ PEL in 1997/1998. The greatest magnitude of change was calculated near

the mouth of the Detroit River, over an expanse of 1,554.75 km² or 6% of the study area. Here sediment contamination was estimated to exceed more than 5.48 µg/g in 1971. Cd concentrations ranging from 3.53 to 10.8 µg/g in 1971 decreased to a range of 0.6 to 1.57 µg/g in 1997/1998, reducing the range of Cd contamination in these sediments from 7.27 µg/g (1971) to 0.96 µg/g (1997/1998). In addition to the reduced Cd levels in the eastern basin by 1997/1998, approximately 24% of the study area reported a downward contamination trend from Cd values ≥PEL, to ≥TEL to <PEL. Between the two survey periods, the majority of the lake (18,593 km²) maintained a consistent level of Cd contamination at the ≥TEL to <PEL interval.

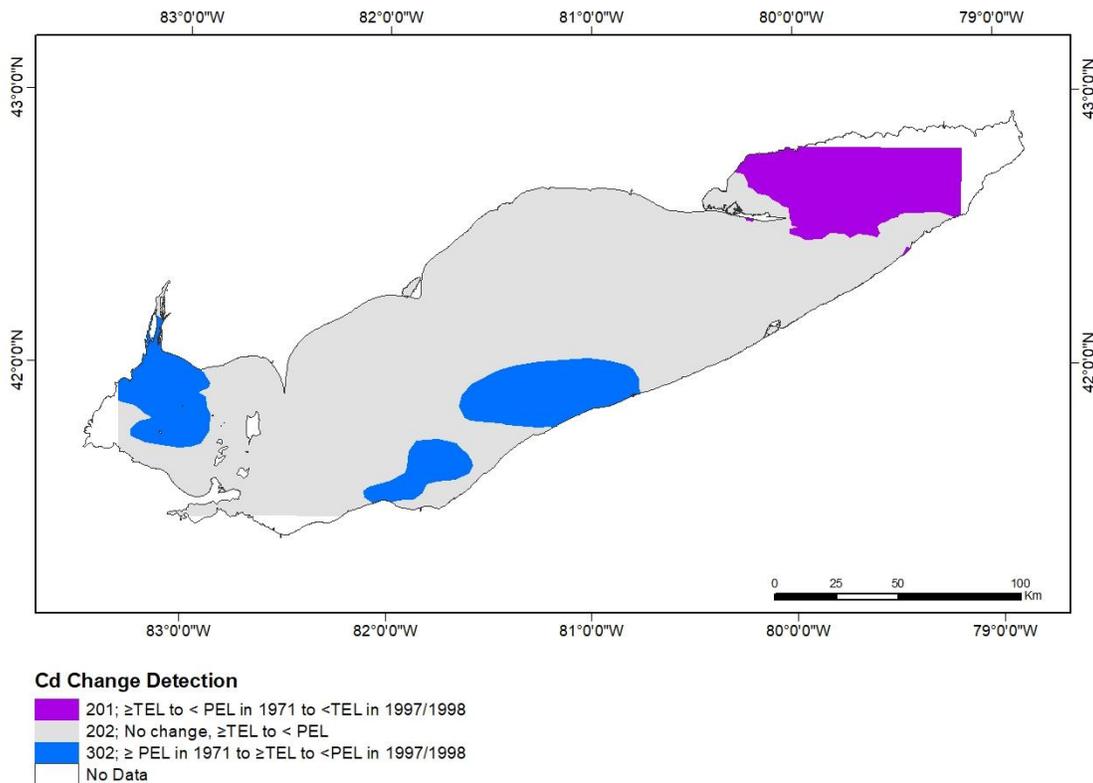


Figure 3-5: Change in Cd contamination concentrations and extent between 1971 and 1997/1998.

Kriging of Cd concentrations and calculating I_{geo} values were completed simultaneously, but are analysed separately. The Raster Calculator was used to calculate the I_{geo} value for the 1971 and 1997/1998 datasets (Esri, 2017). The I_{geo} layer was classified by the Sediment Accumulation Index (Müller, 1981) with intervals ranging from <0 = “practically unpolluted” to >5 = “very strong pollution” (Table 3-3). The same method of change detection was applied to the I_{geo} maps from both years, just as it was employed with the kriged maps. The classified I_{geo} maps for 1971 and

1997/1998 were analyzed with the interpolated concentration maps, and the substrate map featured in Figure 3-6. Excessive sediment contamination from anthropogenic sources was revealed throughout Lake Erie upon accounting for the naturally occurring levels of Cd <1 µg/g. Patterns of contamination <TEL and ≥PEL are strongly reflected by the distribution of I_{geo} accumulation indices; classes of geoaccumulation do provide some context to Cd contamination ≥TEL to <PEL, representing a broad range of contamination with uncertain implications.

From the 1971 survey, sediment along the Long Point-Erie sill was determined to be “practically uncontaminated” ($I_{geo} = 0$) by Cd aside from that occurring naturally. This level of pollution intensity was also identified in a section of the eastern basin, and along the northwestern shore of the central basin. Contamination levels ≥TEL to <PEL ranging from 1.58 – 3.53 µg/g were classified as “unpolluted to moderately polluted” ($I_{geo} = 1$) by the sediment accumulation index. I_{geo} indices calculated “moderate pollution” (5,573.25 km²) that exceeded the isoline boundaries of Cd concentrations predicted to be ≥PEL (3,246 km²) in 1971. The highest values of anthropogenic Cd were identified in sediments at the mouth of Swan Creek, MI, southwest of the Detroit River inlet to Lake Erie. Cd concentrations were predicted to exceed 5.48 µg/g from the 1971 survey; a subset of this highly contaminated sediment was identified to be “moderately to strongly polluted” from anthropogenic sources of Cd alone.

Where Cd sediment contamination was predicted to be <TEL in the 1997/1998 survey, the I_{geo} index also measured the levels of contamination in this region to be “practically uncontaminated” (Figure 3-7). This degree of pollution intensity suggests that traces of Cd in eastern basin sediments are naturally occurring. Along the southern shore of the central basin, most Cd values predicted between 1.58 and <2.55 µg/g were measured to contain “unpolluted to moderately polluted” sediments on the I_{geo} index. Several small areas of “unpolluted to moderately polluted” sediment stand out against the “practically uncontaminated” state of the western basin; most notably the approximate 36 km² area near the mouth of Swan Creek.

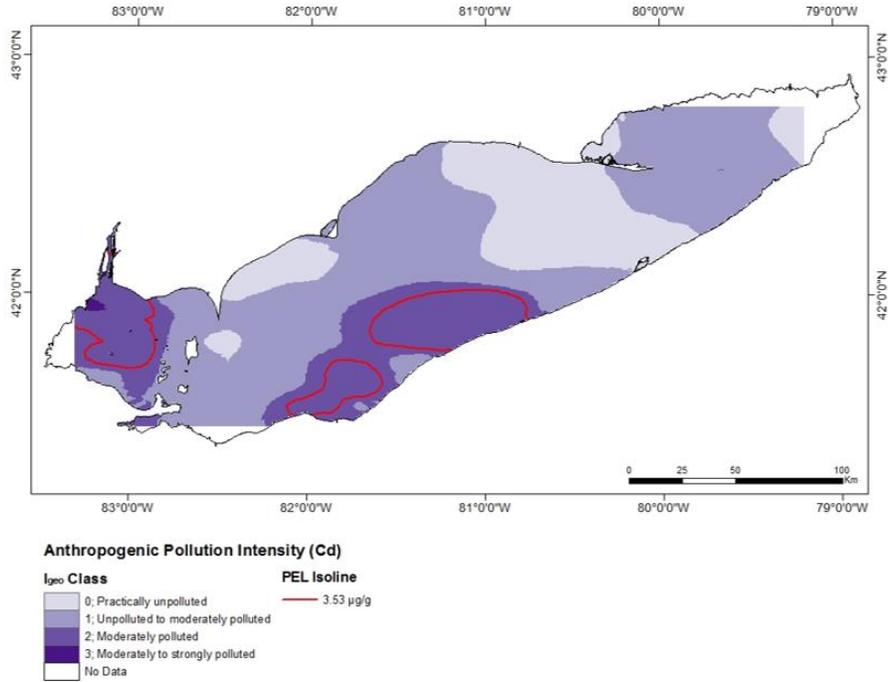


Figure 3-6: Anthropogenic pollution intensity measured by the index of geoaccumulation in contrast to the PEL isoline for the 1971 Cd dataset.

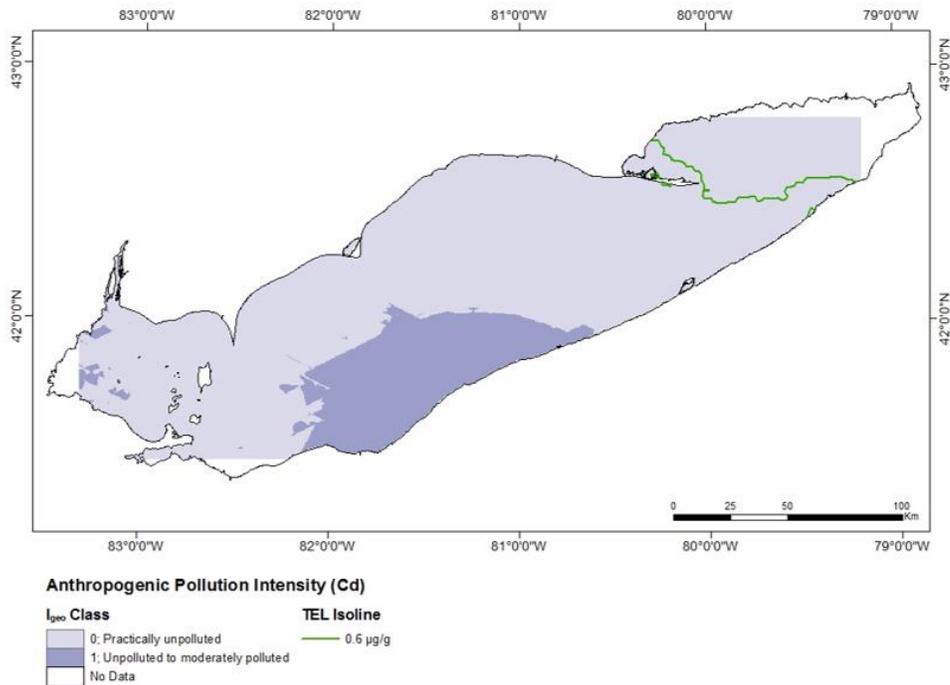


Figure 3-7: Anthropogenic pollution intensity measured by the index of geoaccumulation in contrast to the PEL isoline for the 1997/1998 Cd dataset.

3.5 Raster Overlay and Sediment Substrate Analysis

To assess the patterns and relationships amongst Cd contamination and sediment substrate types, both the 1971 and 1997/1998 predicted contamination maps were reclassified to <TEL, ≥TEL to <PEL, and ≥PEL contamination intervals. When overlaid with sediment substrate types, strong patterns of Cd contamination appear in regions of the lake dominated by mud (clay and silt), and glacial till.

In 1971, contamination ≥PEL represented 3,246 km² or 13% of the study area (Figure 3-8). When overlaid with the substrate map, 90.44% of contamination ≥PEL was identified in regions dominated by mud (68.86%) and sand/mud (21.58%). The remaining 9.56% of the study area with contamination ≥PEL was identified in regions of bedrock (3.74%), sand/gravel (3.54%), and glacial till (1.66%). The majority of Cd contamination in 1971 was predicted to range from ≥TEL to <PEL, representing 21,227.75 km² or 87% of the study area; presence of contamination values <TEL were not identified from the 1971 survey.

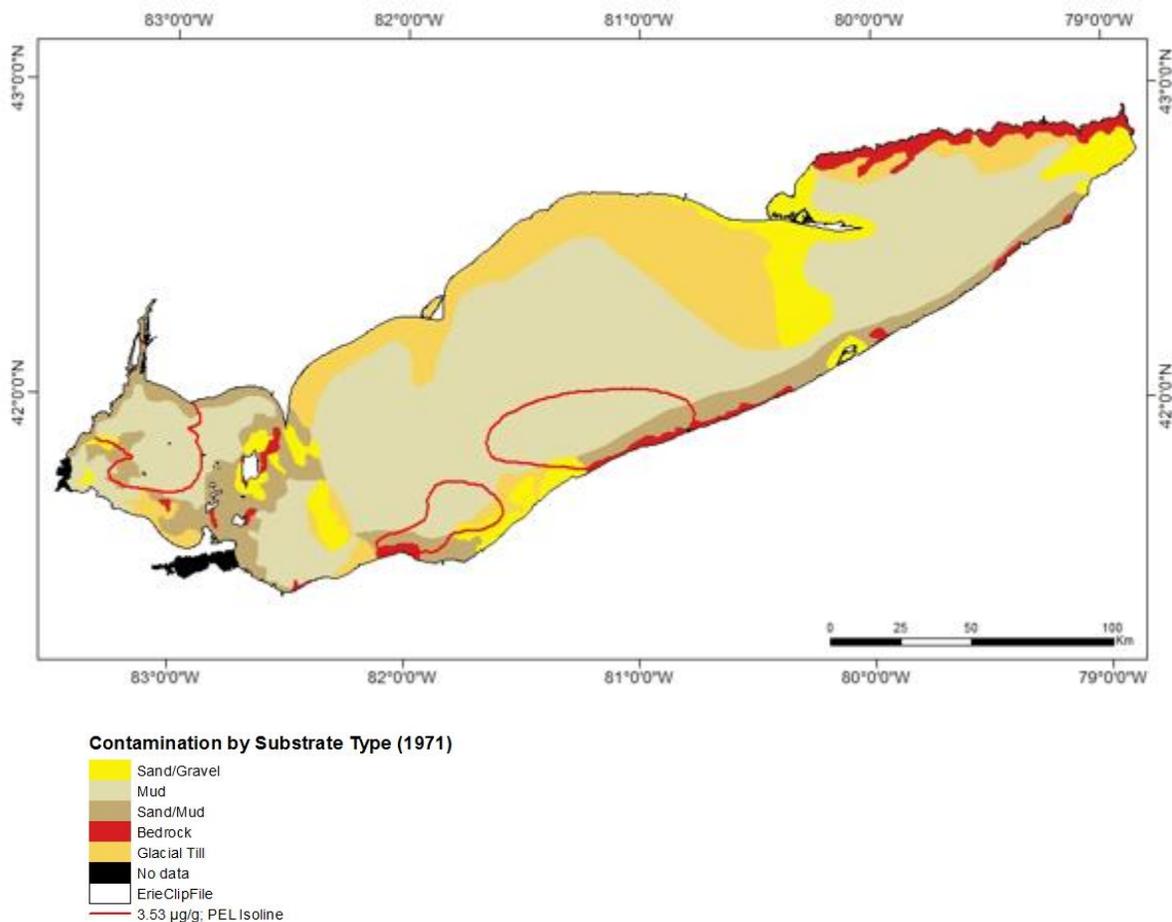


Figure 3-8: Cd contamination ≥TEL to <PEL and ≥PEL from the 1971 dataset by substrate type.

Again, most of contamination values \geq TEL to $<$ PEL were identified in substrate regions dominated by mud (56.55%), as well as glacial till (20.85%). The remaining contamination \geq TEL to $<$ PEL was identified in sand/mud (10.96%), sand/gravel (9.28%), bedrock (1.9%), and no data ($>$ 1%). Surface area (km²) coverage of Cd contamination \geq TEL to $<$ PEL and \geq PEL by substrate type are presented in Table 3-10.

Table 3-10: Surface area (km²) coverage of Cd contamination \geq TEL to $<$ PEL and \geq PEL by substrate type from the 1971 dataset.

Substrate Type	$<$ TEL (km ²)	\geq TEL to $<$ PEL (km ²)	\geq PEL (km ²)
Mud		12,033.25	2,021.5
Sand/Mud		2,331	633.5
Sand/Gravel	n/a	1,975	115
Glacial Till		4,435.5	54
Bedrock		404	121.25

Similar relationships were identified between Cd contamination, mud, and glacial till from the 1997/1998 survey (Figure 3-9). In 1997/1998, contamination \geq TEL to $<$ PEL represented 21,937.5 km² or 89% of the study area; presence of contamination values \geq PEL were not identified from the 1997/1998 survey. The majority of Cd contamination \geq TEL to $<$ PEL was identified in substrate regions of mud (56.71%) while the same range of contamination was found in glacial till covering 18.74% of the study area. The remaining contamination \geq TEL to $<$ PEL was identified in sand and mud (13.63%), sand and gravel (9.18%), bedrock (1.39%), and no data ($>$ 1%).

Almost exclusively in the eastern basin, Cd contamination $<$ TEL represents 2,581 km², or 11% of the study area. The majority (70.75% of the study area) of sediments contaminated with Cd $<$ TEL were identified in substrate regions of mud. Low levels of contamination were also found in regions dominated by glacial till (15.57%), as well as bedrock (8.6%), sand and gravel (3.12%), sand and mud (1.85%), and no data ($>$ 1%). Surface area (km²) coverage of Cd contamination and $<$ TEL and \geq TEL to $<$ PEL by substrate type are presented in Table 3-11.

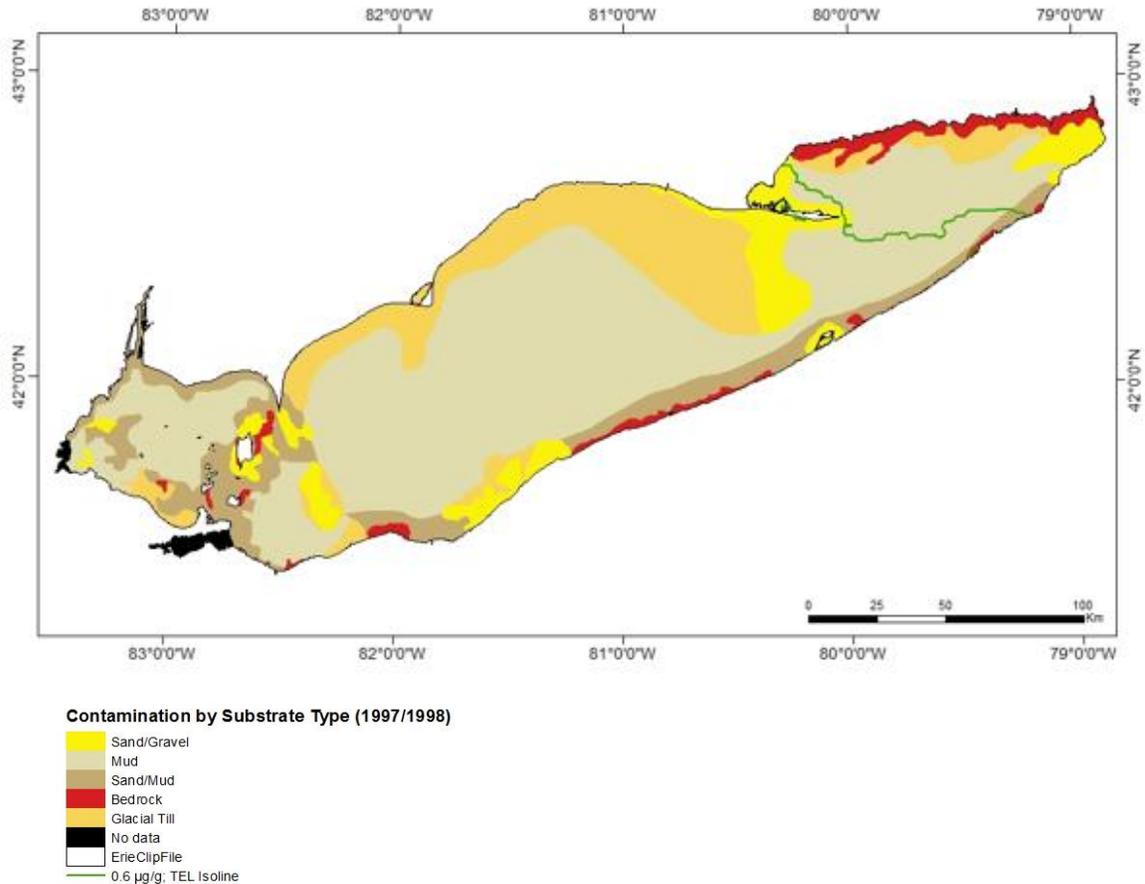


Figure 3-9: Cd contamination <TEL and ≥TEL to <PEL from the 1997/1998 dataset by substrate type.

Table 3-11: Surface area (km²) coverage of Cd contamination <TEL and ≥TEL to <PEL by substrate type from the 1971 dataset.

Substrate Type	<TEL (km ²)	≥TEL to <PEL (km ²)	≥PEL
Mud	1,826	12,440.50	
Sand/Mud	47.75	2,990.50	
Sand/Gravel	80.5	2,014.5	n/a
Glacial Till	401.75	4,111.75	
Bedrock	222	304.25	

3.6 Discussion and Conclusion

Prior to the enactment of the Great Lakes Water Quality Agreement (GLWQA), the Great Lakes Steel Co., Allied Chemical Corp., Ford Motor Co., and Scott Paper Co. were identified as sources of industrial, inorganic and organic pollution in the Detroit River, Huron River and Swan Creek (USEPA, 1965). Until 1967, spoils from the dredging of navigational routes in the Detroit River were often dumped into the open waters of Lake Erie (USAED, 1974); the impacts of dredging

activities were acknowledged for their economic benefits regardless of environmental degradation (Miller, 1998). At the mouth of the Swan Creek, a small section of Lake Erie was identified to have Cd contamination \geq PEL in 1971, and a higher I_{geo} Index value than surrounding areas in 1997/1998 (Figure 3-10). These research findings align with previous studies which used enrichment factors to establish connections between patterns of sediment contamination and historic anthropogenic activities in the lower Great Lakes (Marvin et al., 2002).

The GLWQA has likely been the driving force in heavy metal pollution reduction in all four American shoreline states (Michigan, Ohio, Pennsylvania and New York) and the Canadian Province of Ontario (IJC, 1972). Since 1972, the GLWQA has strived to “restore and maintain the chemical, physical, and biological integrity of the Great Lakes Basin Ecosystem” (IJC, 2002). Interpolated sediment contamination maps show a lake-wide trend of decreasing Cd levels \geq PEL from 1971 to 1997/1998. By 1997/1998, 89.4% of the interpolated area was predicted to contain Cd contamination levels between the \geq TEL to $<$ PEL threshold; sediment in the remaining area (almost exclusively in the eastern basin) was measured to be contaminated with Cd $<$ TEL, and below the background levels of naturally occurring Cd. Specifically, the GLWQA is a form of administrative law, which intends to alter human behavior towards environmental stewardship and responsibility in the Great Lakes Basin (IJC, 1982; IJC, 1994; Tran, 2009), another potential factor leading to the overall lake-wide decline in Cd contamination.

Visualizing Cd contamination patterns by the CCME threshold levels identified regions of the lake where adverse environmental effects were likely to pose health risks to humans and aquatic life. When compared against the Geoaccumulation Index, high-risk regions (i.e. \geq PEL) fell within the boundaries of anthropogenically sourced Cd contamination. Since natural background levels of Cd in Lake Erie are measured at $<1 \mu\text{g/g}$, the \geq TEL to $<$ PEL threshold of contamination is subject to interpretation of Cd values ranging from $>0.6 \mu\text{g/g}$ to $<1 \mu\text{g/g}$. Contamination values in this range are considered to be naturally occurring traces of Cd, and are not likely to pose adverse biological effects, being so close to the TEL. That being said, Cd values above $1 \mu\text{g/g}$ and $<$ PEL are increasingly likely (up to $<3.53 \mu\text{g/g}$) to cause adverse biological impacts from anthropogenic sources (I_{geo} values >1).



Figure 3-10: Extent of anthropogenic contamination identified by the I_{geo} Index around the tributaries of Swan Creek and the Huron River in the western basin of Lake Erie. I_{geo} class = 3, 1971, outlined in dark purple; I_{geo} class = 1, 1997/1998, outlined in light purple (Esri, 2017).

All Cd contamination identified to be \geq PEL in 1971 is also suggested to be of anthropogenic origins by the I_{geo} Index. Overlay analysis identified these regions of high contamination to be in substrates dominated by mud. Moderate levels of Cd contamination ranging from \geq TEL to $<$ PEL in 1971 were predominately located in glacial till substrates. Since glacial till is a medium for naturally occurring Cd, it was not surprising to identify contamination values ranging from 0.6 to $<$ 1 $\mu\text{g/g}$ along the Long Point-Erie sill, and northwestern shoreline of Lake Erie, upon further investigation. Anthropogenically sourced contamination in 1997/1998 was identified by the I_{geo} Index along the Ohio shorelines outside of the traditionally industrious city of Cleveland, Ohio whereas the eastern basin of Lake Erie measured contamination $<$ TEL and as “practically uncontaminated” by the I_{geo} Index. The regression analysis conducted prior to undertaking the kriging process identified a negative correlation between Cd values and depth, where high Cd values were found at shallow depths throughout the lake.

The fundamental multivariate element of co-kriging proved to be valuable in evaluating Cd contamination in lake sediments to be of anthropogenic origins, or naturally occurring. By integrating sediment substrate data into a multivariate raster overlay analysis, interpolated Cd

contamination surfaces were provided with additional context to the distribution of contaminated sediments between 1971 and 1997/1998, and effectively assess spatiotemporal changes in sources of sediment contamination throughout Lake Erie. Two overlapping priorities became clear when Cd contamination by substrate type was displayed by CCME intervals, and the I_{geo} Index; levels of Cd contamination in fine-grained sediments which pose the greatest risk of adverse biological impacts (\geq PEL) are likely the result of direct pollution from anthropogenic sources. This culmination of indicators heavily weighs in favour of long-term sediment remediation efforts and environmental policy implementation specifically in the western Lake Erie basin and along the Ohio shoreline.

Chapter 4: Multivariable 3D Geovisualization of Historic and Contemporary Lead Sediment Contamination in Lake Erie

Abstract: Lead sediment contamination in Lake Erie stems from a long history of natural and synthetic resource production. Historic and contemporary sediment samples were collected by the Canada Centre for Inland Waters in 1971, 1997/1998, and 2014. Each survey is composed of different sampling densities due to the rising cost of sampling procedures and resources. To draw comparisons between sediment surveys over time and space, the kriging interpolation method was used to create continuous data surfaces of lead sediment contamination for each survey. Change detection analyses identified an overall decreasing trend in high levels of lead sediment contamination from 1971 (36.72%) to 1997/1998 (0%) to 2014 (0.71%). Moderate contamination levels remained relatively consistent in surface area coverage between the three survey periods (47.14% in 1971, 51.18% in 1997/1998, and 50.91% in 2014). Sediments with the lowest concentrations of lead contamination increased in surface area throughout Lake Erie coincident with the banning of leaded gasoline in Canada and the United States in the 1990s (12.98% in 1971, 43.85% in 1997/1998, and 38.62% in 2014). Overall lake contamination changed from a basin-wide concern of high lead sediment contamination to a localized situation near Cleveland, Ohio. Lake-wide circulation patterns and bathymetric data were added to interpolated contamination surfaces to enhance the understanding of interrelated hydrodynamic process and geophysical features in the movement of contaminated sediments. Using advanced visualization tools in Esri's ArcScene, bathymetric data were employed to enhance the geographic context of contamination maps. Highly contaminated sediments appear to follow circulation patterns from the mouth of the Detroit River, along the Ohio shoreline where currents are strongest throughout the lake. The physical barriers to sediment transportation created by the Pelee-Lorraine ridge and the Long Point-Erie sill can be visualized in three-dimensions. These elevated features between lake basins are easily recognized as impedances to lake currents when circulation directions are draped over the bathymetric model. By using illumination tools and techniques, geovisualizations of lead sediment contamination throughout Lake Erie create a scientific communication tool for a wide audience to use in multiple-criteria decision making for environmental remediation of sediment contamination.

Key Words: Lead, Kriging, 3D Geovisualization, Great Lakes, Sediment Contamination, Geostatistics, Spatial Analysis

4.1 Introduction

The Great Lakes basin is an industrial, manufacturing, and agricultural powerhouse with a lengthy history of resource production. As one of the world's largest interconnected freshwater systems, the Great Lakes also provide fresh drinking water to over 30 million Canadians and Americans (USEPA, 2015). Lake Erie is the smallest Great Lake by volume at 484 km³ (GLIN, 2016). It is surrounded by intensive farming activities, and several historically industrial cities (Toledo, Ohio; Cleveland, Ohio; Erie, Pennsylvania; and Buffalo, New York; Figure 4-1). Organic and inorganic contamination of Lake Erie has been an environmental concern since the mid-19th century (Evers et al., 2011). Toxic heavy metals, including lead (Pb), pose long-term risks of environmental contamination, as well as bioaccumulation in aquatic species (Wuana and Okieimen, 2011; Khosh Egnhbal, 2014). When ingested by humans, Pb presents critical behavioral, reproductive, and physiological health problems (Kot-Wasik, 2005). Since 1990, both Canada and the United States have banned the use of leaded gasoline under the Canadian Environmental Protection Act (Stephanson, 2009) and the Clean Air Act (USEPA, 1996) resulting in fewer cases of human Pb poisoning (Kot-Wasik, 2005).



Figure 4-1: Lake Erie and its basins (Marvin et al., 2002; ESRI, 2017).

Hydrodynamic conditions and bathymetric contours influence the spatial patterns of Pb sediment contamination throughout Lake Erie. Until recently, visualizations of sediment contamination (Forsythe et al., 2004) and internal lake processes (Beletsky et al., 1999) were presented in two-

dimensions (2D). In 2014, the National Oceanic and Atmospheric Administration (NOAA) introduced a complete bathymetric model of Lake Erie and its basin (NOAA, 2014; Figure 4-2). Hundreds of thousands of soundings measurements collected for over 100 years by the U.S. Army Corps of Engineers, the NOAA Coast Survey, and the Canadian Hydrographic Service were compiled to create the *Bathymetry of Lake Erie and Lake Saint Clair* (NOAA, 2014).

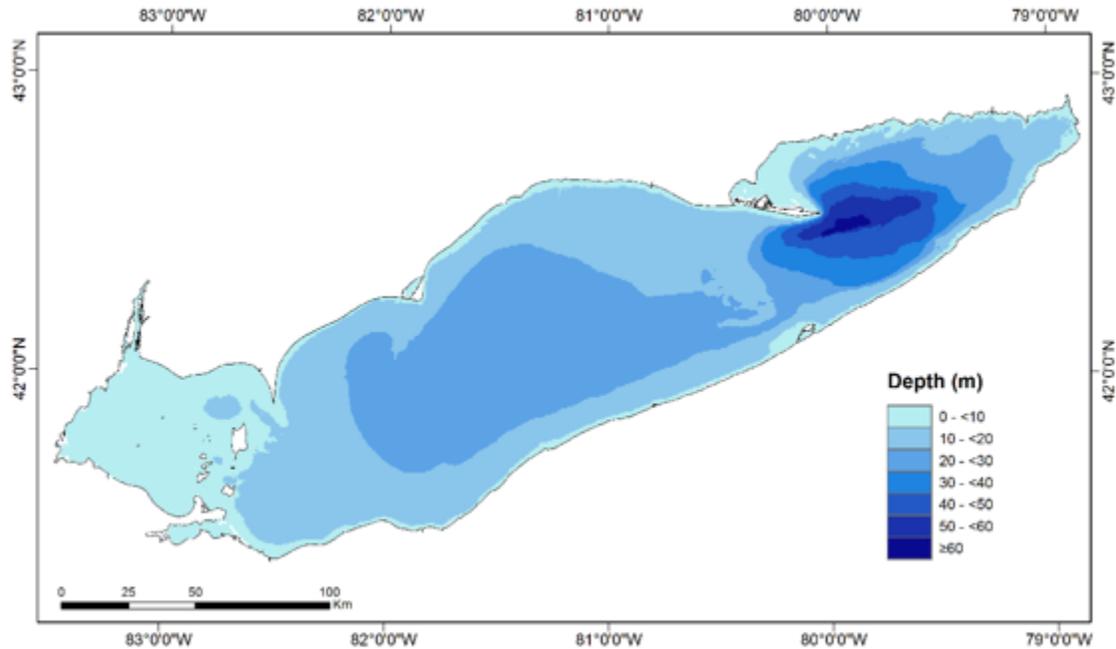


Figure 4-2: Lake Erie depth (m) map (NOAA, 2014).

The innovative bathymetric model allows for the dynamic relationship between sediment contamination and lake circulation patterns to be visualized in three-dimensions (3D). This research intends to produce novel 3D geographic visualizations of Pb sediment contamination from 1971, 1997/1998, and 2014 for two purposes; to increase the spatial awareness of sediment contamination patterns over time, and to create approachable yet valuable pollution maps to be interpreted by a wide variety of users for remediation decision making.

4.1.1 Pb Contamination in Lake Erie

Environmental Pb contamination stems primarily from anthropogenic sources including agricultural pesticide and fertilizer use (Wuana and Okieimen, 2011), industrial manufacturing of batteries, and air pollution from fossil fuel combustion (Kot-Wasik, 2005; Wuana and Okieimen, 2011). Anthropogenic Pb pollution is known to have a strong affinity for clay sediments (Soong, 1974; Jain and Ram, 1997). Naturally occurring Pb is minimal at background levels of 18.2 $\mu\text{g/g}$

(Marvin et al., 2004; ATSDR, 2007). Trace levels in the natural environment originate from volcanic eruptions and erosion of crustal rock containing Pb minerals (ATSDR, 2007).

4.1.2 Lake Erie Circulation

Hydrodynamic processes underlie most lake-wide physical and biological activities (Bai et al., 2013). Waves, currents and circulation patterns influence the distribution and displacement of sediments, as well as the rate at which contamination is transported throughout the basin (Beletsky et al., 1999; Rao and Schwab, 2007; Bai et al., 2013). Episodic waves created during storm events have a greater impact on the movement of sediments in the nearshore coastal zone (0-20 m deep, Figure 4-2) than average climatic conditions (ERSPA, 2015) and circulation throughout the lake (Rao and Schwab, 2007; Morang et al., 2011). A rise in major storms and weather extremes resulting from climate change are anticipated to impact the biophysical characteristics and hydrodynamic processes of the Great Lakes basin (Blais and Kalff, 1995; CCME, 2003; Morang et al., 2011; Verma et al., 2015). For example, less lake-wide ice coverage throughout the winter months may lengthen the lakes storm season, ultimately exposing vulnerable coastal regions to more frequent opportunities for sediment resuspension (ERSPA, 2015).

Prevailing westerly winds, parallel to the basin's orientation, are largely responsible for the annual circulation patterns of Lake Erie (Beletsky et al., 1999; Bai et al., 2013; Figure 4-3). This strong, consistent wind creates a counter rotating 2-gyre circulation system within the relatively flat bathymetry of the central basin (Fujisaki et al., 2012; Bai et al., 2013). Lake Erie is the shallowest Great Lake averaging 19 m deep, reaching a maximum of 64 m in the eastern basin (GLIN, 2016). Strong currents flow along the shallow shorelines with the prevailing westerly winds (Fujisaki et al., 2012; Bai et al., 2013) reaching upwards of 3.7 cm/s along the Ohio shoreline, near Cleveland (Beletsky et al., 1999). Pressure gradients drive the westward flow in deep offshore regions of the central basin (Beletsky et al., 1999) creating an anticyclonic gyre (Figure 4-3a) along the Ontario shoreline, and a cyclonic gyre (Figure 4-3b) along the Ohio shoreline.

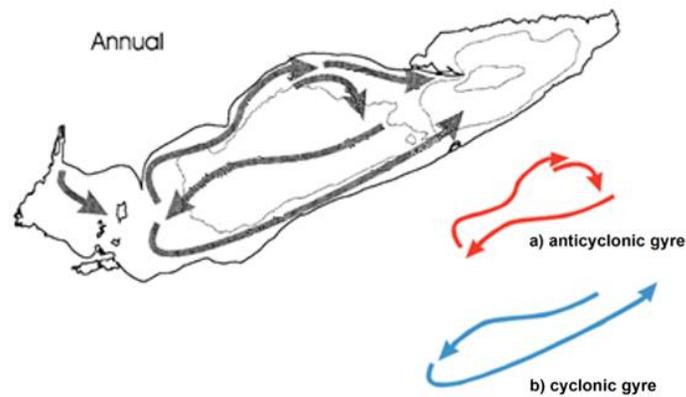


Figure 4-3: Annual Lake Erie circulation patterns with 20m and 50m bathymetry isolines (Beletsky et al., 1999); the two-gyre system in the central basin is identified individually by the a) anticyclonic gyre and b) cyclonic gyre.

Stronger wind-driven currents and waves can develop over the central basin due to the flat and shallow nature of the basin bathymetry (Pincus, 1953; Beletsky, 1999). With few physical impedances to reduce wind stresses, the central basin is vulnerable to frequent sediment resuspension (Alther, 1981; Hawley and Eadie, 2007; Morang et al., 2011). At a maximum depth of 23.5 m, the central basin is almost entirely classified within the coastal zone (<20 m), where wind-driven wave activity fosters sediment resuspension and transportation (Alther, 1981; ERSPA, 2015). The shorelines of Lake Erie are also defined by very steep slopes (>45 degrees) and shallow littoral zones (Duarte and Kalf, 1986; Derosier et al., 2015). Fine grained sediment substrates most vulnerable to erosion and transportation (i.e. clay, silt and sand) dominate the southern shoreline where wind-driven currents are strongest (Beletsky, 1999). Sediments in the littoral zone are more likely to be resuspended and transported by wave action (Derosier et al., 2015) especially on steep slopes where unconsolidated sediments are vulnerable to gravitational and erosional processes (Christopherson and Birkeland, 2015).

4.1.3 Geovisualization

The term “geovisualization” is a product of *Visualization in Scientific Computing* by McCormick et al. (1987). Geovisualization is a method of displaying geographic data, or thematic information as a visual tool for exploratory data analysis (EDA), knowledge construction and decision making (Andrienko and Andrienko, 2003; MacEachren et al., 2004; Koua et al., 2006). When employed under data analysis conditions other than EDA, Rinner (2007) suggests the term “visual analytics”

as a more appropriate name for the visualization of processed data. Regardless of the stage in which geovisualizations are employed, Andrienko and Andrienko (2003), and Ruda (2016) recommend its integration with multiple-criteria decision making (MCDM).

Visualizing geographic data as maps or graphs illustrates spatial trends, patterns and relationships between multivariate and multidimensional geographic processes and phenomena (Koua et al., 2006, Rinner, 2007; Kumar et al., 2013). When available, digital elevation models (DEM) and bathymetric data can enhance the perception of spatial data patterns within their geographic context (Rinner, 2007; Lloyd and Dykes, 2011; Knust and Buchroithner, 2014). This degree of perception (especially of geometric shapes and contours) can be illuminated by 3D visualization tools (i.e. hillshade) and lighting effects (Knust and Buchroithner, 2014). By effectively draping thematic data layers (of sediment contamination, for example) onto 3D geographic contours, complex multivariate data analyses are possible through a novel, yet approachable illustration of spatial data interacting with underlying geomorphological processes (Koua et al., 2006; Lloyd and Dykes, 2011).

4.2 Data

The Great Lakes Basin Sediment Database (Environment and Climate Change Canada) houses sediment surveys conducted by the Water Science and Technology Directorate (previously the National Water Research Institute) and the Great Lakes Surface Water Surveillance Program (Marvin et al., 2002; Rukavina et al., 2013) since 1965. Sediment contamination in Lake Erie was first sampled in 1971 (263 survey locations), and again in 1997/1998 (55 survey locations) and 2014 (34 survey locations). The drastic change in survey density was the result of increasing procedural costs between the three sampling periods (Forsythe et al., 2010). Descriptive statistics for each sediment survey are presented in Table 4-1.

Table 4-1: Descriptive statistics of Pb sediment contamination samples from Lake Erie.

Year	Min. (µg/g)	Max. (µg/g)	Mean (µg/g)	Standard Dev.	Variance
1971	9.30	299.30	85.98	49.55	2,455.20
1997/1998	4.98	104.27	43.08	23.21	538.70
2014	5.00	168.00	40.61	27.49	755.70

4.2.1 Kriging

Kriging is a geostatistical method of spatial interpolation which is used to estimate contamination values at unsampled locations from existing sample measurements (Ouyang et al., 2003;

Forsythe and Marvin, 2009). This tool will be utilized to transform discrete Pb contamination measures into a continuous data layer. In doing so, sediment surveys of different point densities, and from different collection periods can be more accurately compared over space and time. Kriging also produces valuable cross-validation statistics describing the validity of a predictive model in accurately estimating contamination values at unsampled locations (McBratney and Webster, 1986; Forsythe and Marvin, 2005; Forsythe et al., 2010).

Conventional standards of a statistically valid, unbiased kriging model include a Mean Prediction Error (MPE) close to 0, an Average Standard Error (ASE) less than 20, a Standardized Root-Mean-Squared Prediction Error (SRMSPE) close to 1, and a minimal difference between the Root-Mean-Squared Prediction Error (RMSPE) and ASE (Johnston et al., 2001; Simpson and Wu, 2014; Forsythe et al., 2016a; Forsythe et al., 2016b; Forsythe et al., 2016c). In addition to estimating values at unsampled locations, kriging interpolation also predicts a value at all sampled locations from which the model is identified to underestimate ($MPE > 0$) or overestimate ($MPE < 0$) existing measurements (Osburn, 2000; Forsythe and Marvin, 2005). Under- or overestimation of variability made by the kriging model is also represented by SRMSPE values greater than or less than 1 $\mu\text{g/g}$, respectively (Johnston et al., 2001; Forsythe et al., 2016b; Forsythe et al., 2016c).

4.2.2 Canadian Sediment Quality Guidelines for the Protection of Aquatic Life (CCME)

The degrees of Pb sediment contamination which threaten the health and biological wellbeing of humans and aquatic life have been established by the Canadian Council of Ministers of the Environment (CCME, 1999; Forsythe and Marvin, 2009). The Threshold Effect Level (TEL) for Pb contamination is 35.00 $\mu\text{g/g}$ (CCME, 1999). Below this level of pollution, biological risks to humans and aquatic life are not likely to occur (Forsythe and Marvin, 2009). The Probable Effect Level (PEL) for Pb contamination is 91.30 $\mu\text{g/g}$ (CCME, 1999). Above this level, biological risks to humans and aquatic life are expected to frequently occur (Forsythe and Marvin, 2009). Sediment contamination with respect to the TEL and PEL should be monitored over time for potential changes in concentrations.

4.3 Methods

4.3.1 Kriging

During preliminary data exploration, Pb samples from 1971, 1997/1998, and 2014 were identified to be abnormally distributed from histogram analyses. Kriging cross-validation errors (RMSPE and ASE) at or exceeding their respective conventional standards of statistical validity (Table 4-

2) also suggested abnormally distributed data. Similar geospatial analyses of sediment contamination (Ouyang et al., 2003; Lui et al., 2006; Gawedzki and Forsythe, 2012) recommend a log-transformation of abnormally distributed data sets to improve the accuracy and statistical validity of a kriging procedures (Table 4-3).

Table 4-2: Kriging cross-validation statistics of the 1971, 1997/1998, and 2014 Pb datasets.

Year	MPE (µg/g)	RMSPE (µg/g)	SRMSPE (µg/g)	ASE (µg/g)	RMSPE-ASE (µg/g)
1971	0.508	31.803	0.919	34.699	-2.896
1997/1998	0.482	19.399	0.969	20.592	-1.193
2014	-0.597	27.107	1.029	23.511	3.596

Table 4-3: Kriging (log-normal) cross-validation statistics of the 1971, 1997/1998, and 2014 Pb datasets.

Year	MPE (µg/g)	RMSPE (µg/g)	SRMSPE (µg/g)	ASE (µg/g)	RMSPE-ASE (µg/g)
1971	0.003	0.189	1.033	0.181	0.008
1997/1998	0.011	0.258	1.062	0.251	0.007
2014	0.008	0.248	1.014	0.238	0.010

Each of the three datasets were best modelled using an exponential distribution, and the parameters listed in Table 4-4. The kriging analysis was set to include a minimum of 1 and a maximum of 5 neighbours as influential points in predicting values at unsampled locations. Selecting the most appropriate distribution model was determined through an experimental assessment of different parameter combinations and their respective cross-validation statistics. All kriging models produced SRMSPE values close to 1, and low (~0) RMSPE and ASE values (Forsythe et al. 2004; Table 4-3).

Table 4-4: Kriging parameters used for modeling the 1971, 1997/1998, and 2014 Pb (log-normal) datasets.

Year	Distribution Model	Major Range	Minor Range	Direction (°)	Lag Distance
1971	Exponential	100,000	50,000	90	30,832
1997/1998	Exponential	100,000	50,000	90	28,552
2014	Exponential	100,000	50,000	90	27,086

4.3.2 3D Geovisualization

Lake Erie bathymetric data were acquired in an ARC ASCII file format from the NOAA National Centers for Environmental Information (NCEI). This DEM contains both terrestrial elevation and bathymetric data of Lake Erie and its basin. First, bathymetric data were extracted from the DEM by the surface extent of Lake Erie. A depth profile chart was created using ArcMap's 3D Analyst (Esri, 2017) for preliminary analysis of lake depth across the basin (Figure 4-4).

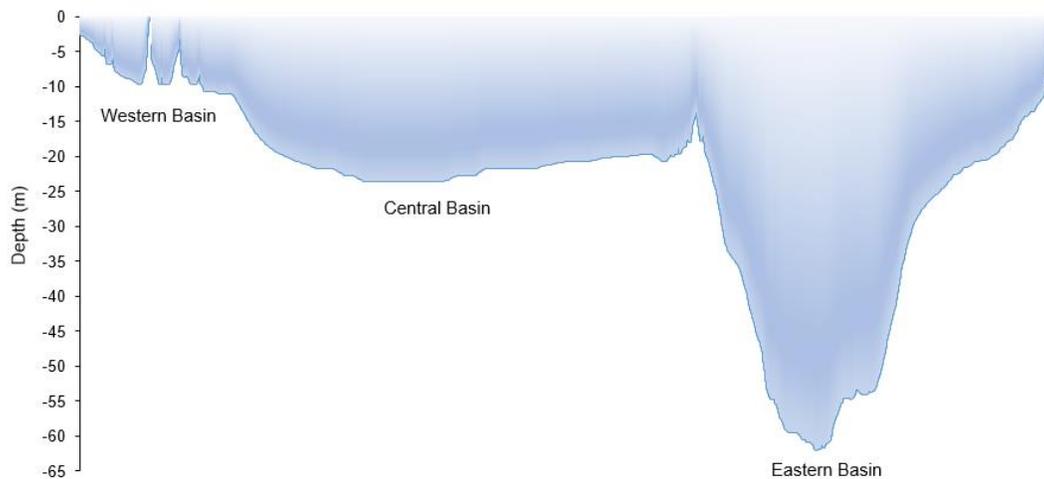


Figure 4-4: Lake Erie depth profile (m).

Esri's 3D viewing application ArcScene (Esri, 2017) was used to create geovisualizations of Pb contamination throughout Lake Erie. This program displays elevation data (Z) from feature geometry and allows for the 3D geovisualization of thematic data layers (3D Analyst, 2017). Average annual lake-wide circulation patterns were digitized (based on maps by Beletsky et al., 1999) and integrated into 3D geovisualizations of Pb sediment contamination patterns. Along with bathymetric data, circulation patterns provide hydrodynamic context to sediment movement throughout the basins of Lake Erie.

The bathymetry of Lake Erie is displayed in Figures 4-5 and 4-6. The crescent-shaped ridges created by the Long Point-Erie sill and the Long Point spit extending into the eastern basin overshadow deeper contours when viewed from the normal southern position (Figure 4-5). The lake viewpoint was therefore rotated to the southeast perspective (Figure 4-6) to best view the complex shape and contours of the eastern basin and to provide the best viewpoint for 3D geovisualization (Forsythe et al., 2016c).

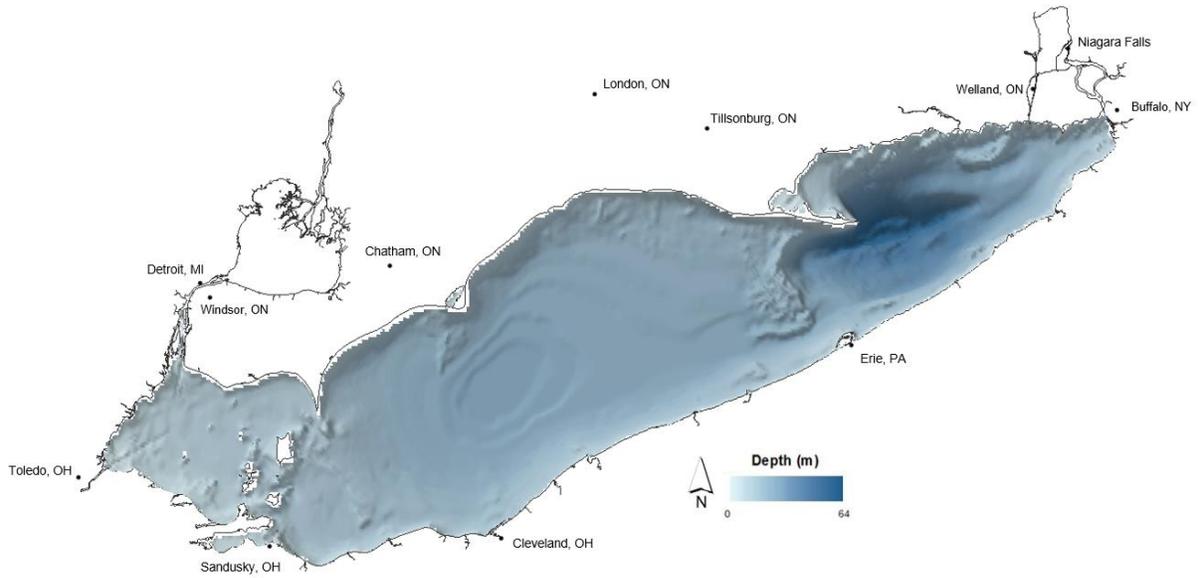


Figure 4-5: Bathymetry of Lake Erie viewed from the normal southern position.

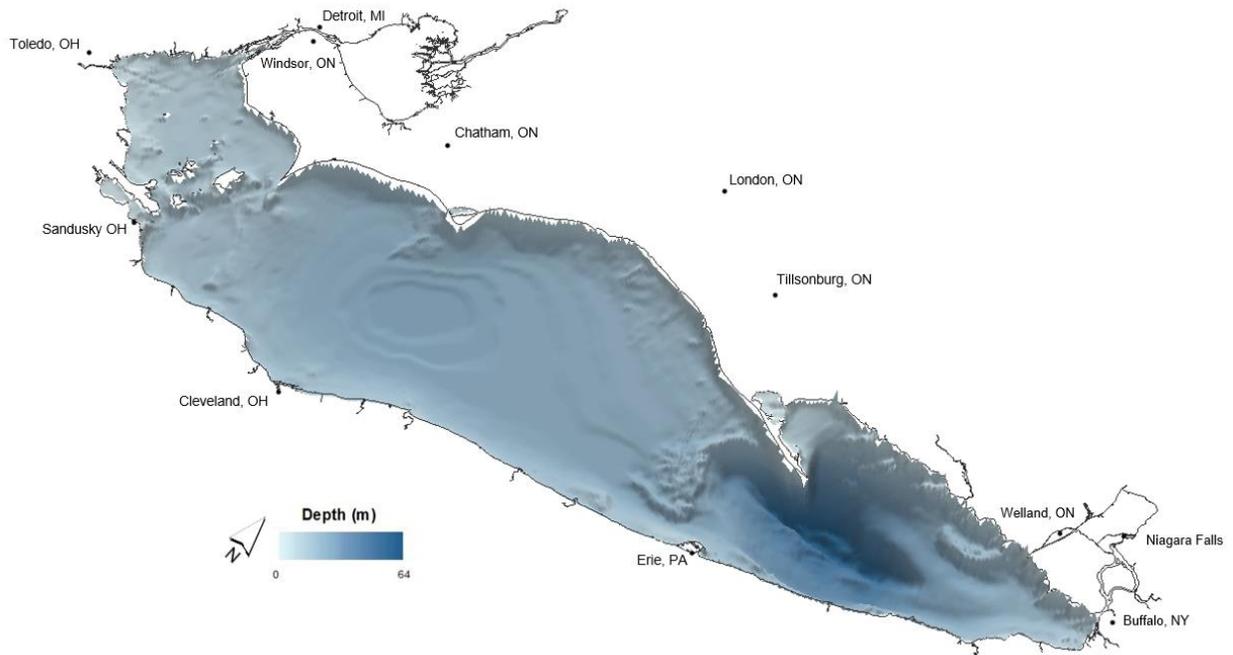


Figure 4-6: Bathymetry of Lake Erie from the southeast perspective.

4.3.3 Change Detection

Change detection analyses were performed between each survey period (1971 and 1997/1998, 1997/1998 and 2014, and 1971 and 2014) to assess the degree and spatial extent of Pb sediment contamination change over time. Contamination intervals in each map were aggregated to <TEL, ≥TEL to <PEL, and ≥PEL categories to simplify the change detection output (Table 4-5). These intervals were then reclassified in each contamination map to unique numeric identifiers, which when added together identified the degree of contamination change between the three survey periods. Each possible degree of change between two Pb contamination maps is displayed and described in Table 4-6.

Table 4-5: Pb (log-normal) contamination intervals <TEL, ≥TEL to <PEL, and ≥PEL.

Contamination Intervals <TEL (µg/g)	Contamination Intervals ≥TEL to <PEL (µg/g)	Contamination Intervals ≥PEL (µg/g)
0.00 - <11.67	≥35.00 - <53.77	≥91.30 - <110.07
≥11.67 - <23.33	≥53.77 - <72.53	≥110.07 - <128.83
≥23.33 - <35.00	≥72.53 - <91.30	≥128.83

Table 4-6: Contamination change categories and their meanings.

Change Category	Description
101	No change, <TEL
102	<TEL to ≥TEL to <PEL
103	<TEL to ≥PEL
201	≥TEL to <PEL to <TEL
202	No change, ≥TEL to <PEL
203	≥TEL to <PEL to ≥PEL
301	≥PEL to <TEL
302	≥PEL to ≥TEL to <PEL
303	No Change, ≥PEL

4.4 Spatial Analysis

4.4.1 Kriging Results

The spatial extent of Pb contamination from 1971, 1997/1998, and 2014 sediment surveys are presented in Table 4-7 and Figures 4-7 to 4-9. In 1971, contamination ≥PEL covered 37.92% of the study area (24,983.5 km²) and was mostly found at the mouth of the Detroit River, and along the shoreline of Ohio following the direction of the strongest annual circulation pattern (Figure 4-7). From this survey, high Pb sediment contamination did not necessarily correlate with depth. Contamination ≥128.83 µg/g was estimated in sediments throughout the shallow western and central basins. In the deepest contours of the eastern basin, the maximum predicted Pb contamination ranged from 91.30 to <128.83 µg/g.

Table 4-7: Spatial extent of predicted Pb (log-normal) contamination by year.

Year	Pb Contamination	<TEL	≥TEL to <PEL	≥PEL	Total Analysis Area	No data
1971	Area (km ²)	3,347.5	12,162.25	9,473.75	24,983.5	816.5
	Area (%)	13.4	48.68	37.92	96.84	3.16
1997/ 1998	Area (km ²)	11,313.5	13,205	n/a	24,518.5	1,281.5
	Area (%)	46.14	53.86	n/a	95.03	4.97
2014	Area (km ²)	9,964.5	13,134.75	182.25	23,281.5	2,518.5
	Area (%)	42.8	56.42	0.78	90.24	10.82

Nearly half (48.68%) of the study area along the northern (Ontario) shoreline, and throughout the central and eastern basins was predicted to have Pb contamination levels ≥TEL to <PEL from the 1971 sediment survey. Sediments surrounding Pelee Island in the western basin and along the Long Point-Erie sill in the central basin were predicted to have the least amount of Pb contamination (13.4%). Pb contamination ranging from <PEL generally aligned with the northeast circulation patterns parallel to the Ontario shoreline from Pelee Island to the Long Point-Erie sill (Alther, 1981). Due to the sediment survey point distribution, some areas of Lake Erie were excluded from the prediction surface. The 1971 dataset allowed for 96.84% of the 25,800 km² lake surface area to be mapped by the kriging prediction model, leaving 816.5 km² of Lake Erie classified as No Data.

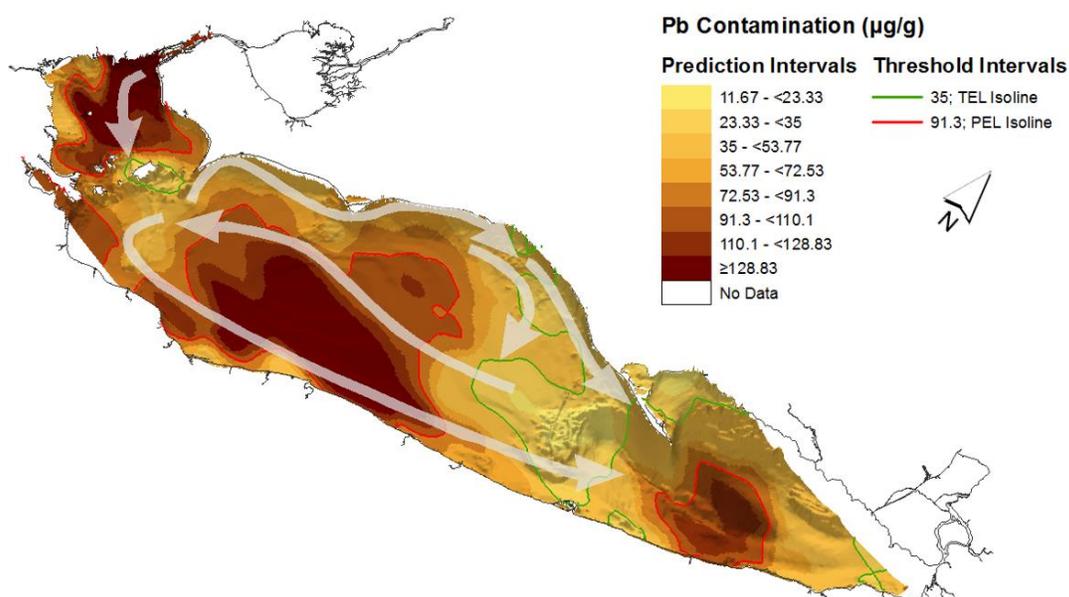


Figure 4-7: Kriging results of the 1971 Pb (log-normal) contamination in Lake Erie with average annual circulation patterns.

Lake Erie sediment contamination did not exceed the PEL in 1997/1998 (Figure 4-8). The TEL isoline roughly divided the lake in half through the central basin; contamination \geq TEL to $<$ PEL to the west and $<$ TEL to the east. The influence of strong eastward circulation along the Ohio shoreline is apparent in the shape of the TEL isoline throughout the central basin. Pb Contamination \geq TEL to $<$ PEL had advanced further towards the eastern basin along the southern shoreline than contaminated sediments to the north. Specifically, 53.86% and 46.14% of Pb contamination from the 1997/1998 survey was predicted to range from \geq TEL to $<$ PEL and $<$ TEL, respectively. Due to the lower point density, and distribution of the 1997/1998 survey samples 95.03% of the study area was included in the kriging analysis.

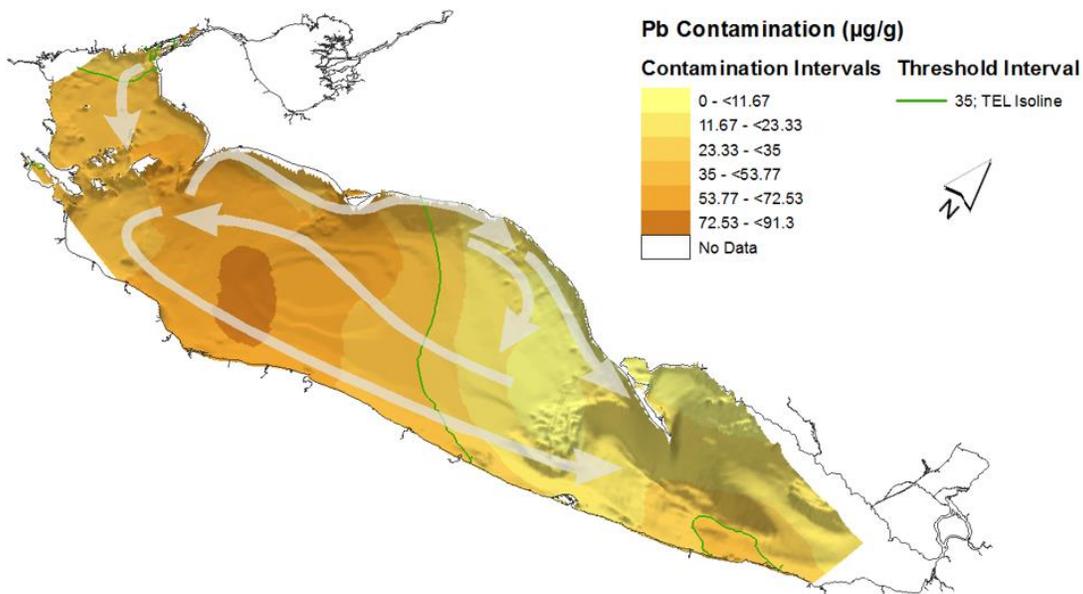


Figure 4-8: Kriging results of the 1997/1998 Pb (log-normal) contamination in Lake Erie with average annual circulation patterns.

The 2014 sediment survey produced contamination prediction patterns similar to those estimated from the 1997/1998 dataset (Figure 4-9). One notable difference is the small portion (0.78% of the study area) of Pb sediment contamination outside of Cleveland. The TEL isoline created a similar division of the lake to that of the 1997/1998 TEL isoline. However, the 2014 dataset estimated slightly more area (56.42%) represented by Pb contamination \geq TEL to $<$ PEL throughout all three basins. Like the 1997/1998 map, contamination \geq TEL to $<$ PEL in the central basin had advanced further eastward along the southern shoreline, following the strong circulation patterns in this region of the lake. Sediment contamination $<$ TEL was predicted to represent 42.8% of the study area predominately in the central and eastern basins along the Ontario shoreline and the

Long Point-Erie sill. The 2014 dataset contained the fewest sampling point locations of the three surveys resulting in the smallest area of analysis at 23,281.5 km² (90.24% of the lake surface area).

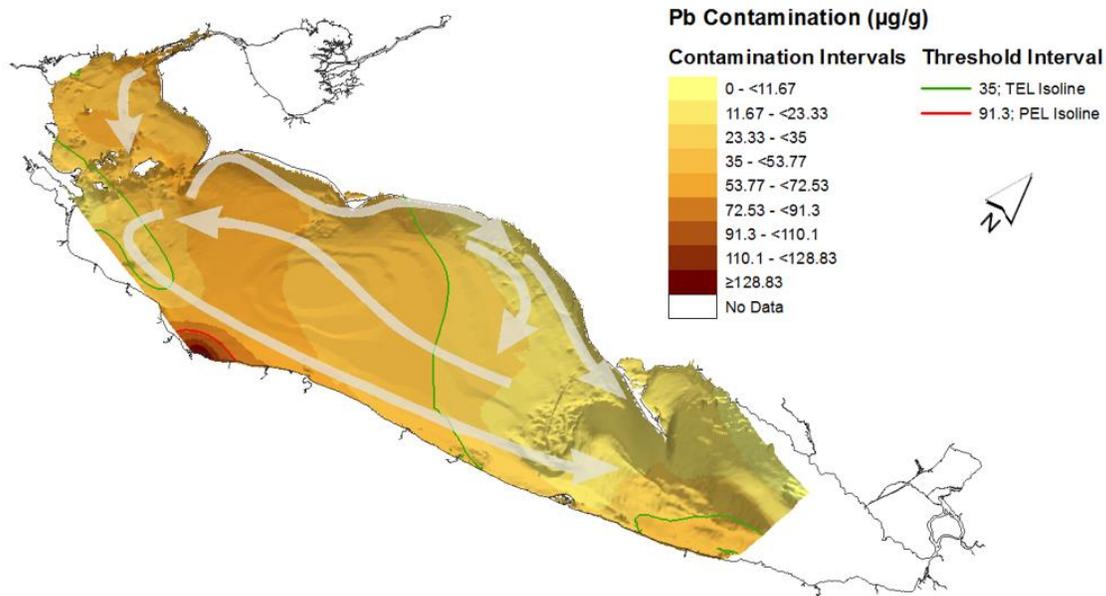


Figure 4-9: Kriging results of the 2014 Pb (log-normal) contamination in Lake Erie with average annual circulation patterns.

4.4.2 Spatiotemporal Change in Pb Sediment Contamination

The degree of change in sediment contamination between 1971 and 1997/1998 is displayed in Figure 4-10. Between both sediment surveys (~27 years apart) more than half (65%) of the lake study area experienced a decline in sediment contamination levels (Table 4-8). Throughout most of the western and central basins (33.41% of the study area), Pb contamination levels \geq PEL in 1971 decreased to \geq TEL to $<$ PEL by 1997/1998. At the mouth of the Detroit River, and in small sections of the central and eastern basins (5.34% of the study area), contaminated sediments \geq PEL in 1971 were drastically reduced to $<$ TEL by 1997/1998. Throughout the eastern half of the central basin and the eastern basin (26.26% of the study area), Pb sediment contamination decreased from \geq TEL to $<$ PEL to $<$ TEL between the two survey periods. The only instance of increased Pb contamination between 1971 and 1997/1998 were in sediments surrounding Pelee Island along the Pelee-Lorraine ridge (0.72% of the study area). The remaining 34.28% of lake study area identified sediments which did not experience any degree of contamination change (based on threshold intervals) between 1971 and 1997/1998. Pb contamination \geq TEL to $<$ PEL remained throughout the western and central basins (21.72% of the study area); along the Ontario

shoreline in the central basin and the Long Point-Erie sill sediment contamination remained at <TEL from 1971 to 1997/1998.

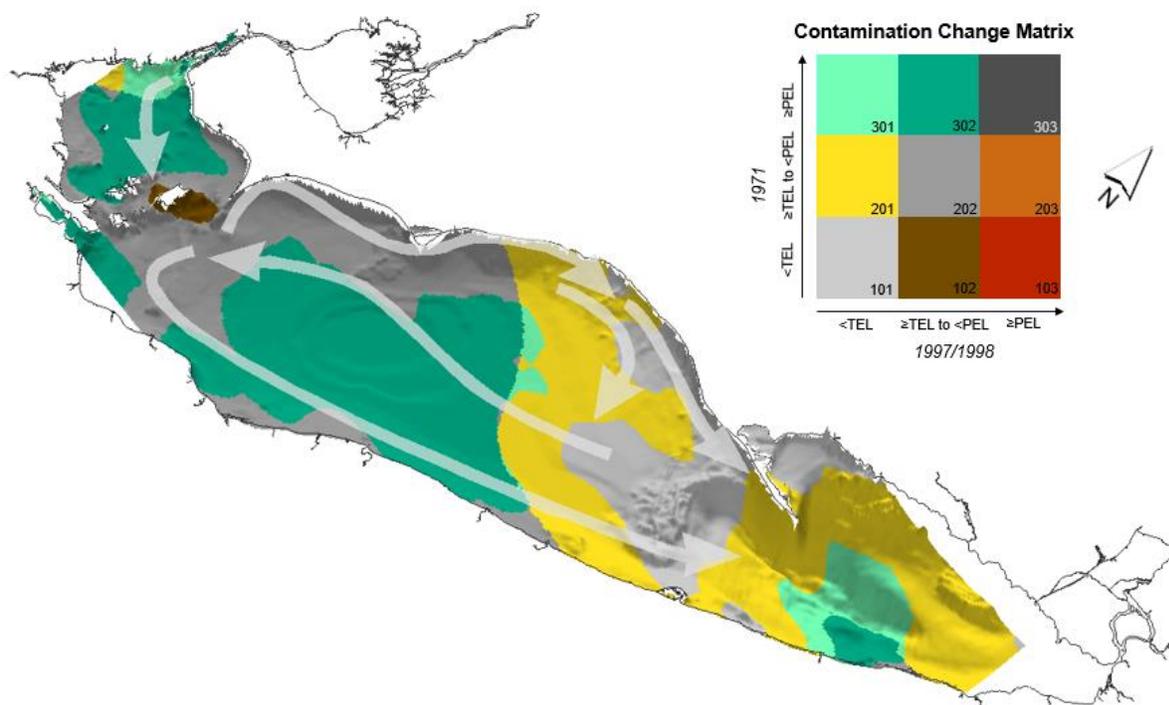


Figure 4-10: Change in Pb (log-normal) sediment contamination between 1971 and 1997/1998.

Table 4-8: Surface area change in Pb (log-normal) sediment contamination between 1971 and 1997/1998.

Change Category	Description	Area (km ²)	Area (%)
101	No Change, <TEL	3,060.5	12.55
102	<TEL to ≥TEL to <PEL	175	0.72
201	≥TEL to < PEL to <TEL	6,404.75	26.26
202	No change, ≥TEL to < PEL	5,297.25	21.72
301	≥ PEL to <TEL	1,302.75	5.34
302	≥ PEL to ≥TEL to <PEL	8,146.5	33.41
No Data		1,413.25	5.8

Concentrations of Pb sediment contamination in Lake Erie remained relatively stable between 1997/1998 and 2014 (Figure 4-11). Contamination measures throughout most of the study area (90.11%) did not change in the 17 years between surveys (Table 4-9). More than half (51.66%) of the sediments maintained contamination values ranging from ≥TEL to <PEL throughout the western basin, and half of the central basin. A distinct boundary runs through the middle of the central basin, where sediment contamination to the east and extending into the eastern basin remained at <TEL between 1997/1998 and 2014 (representing 38.44% of the study area). Only

9.89% of the lake study area experienced change in Pb sediment contamination between the two surveys. A small area (182 km²) outside of Cleveland saw an increase in Pb sediment contamination values, ranging from \geq TEL to $<$ PEL in 1997/1998 to \geq PEL by 2014. Sediment contamination also worsened at the mouth of the Detroit River, and in small sections of the central and eastern basins where Pb measured $<$ TEL in 1997/1998 up to \geq TEL to $<$ PEL in 2014. This change in the central basin suggests an eastward expansion of contaminated sediments ranging from \geq TEL to $<$ PEL. This contamination threshold boundary also identifies a small proportion of contaminated sediments closest to the Ontario shoreline which have improved between the two surveys. Here, as well as along the Ohio shoreline in the western and central basins, Pb contamination values have fallen from \geq TEL to $<$ PEL in 1997/1998 to $<$ TEL in 2014 (representing a total of 1,023 km²).

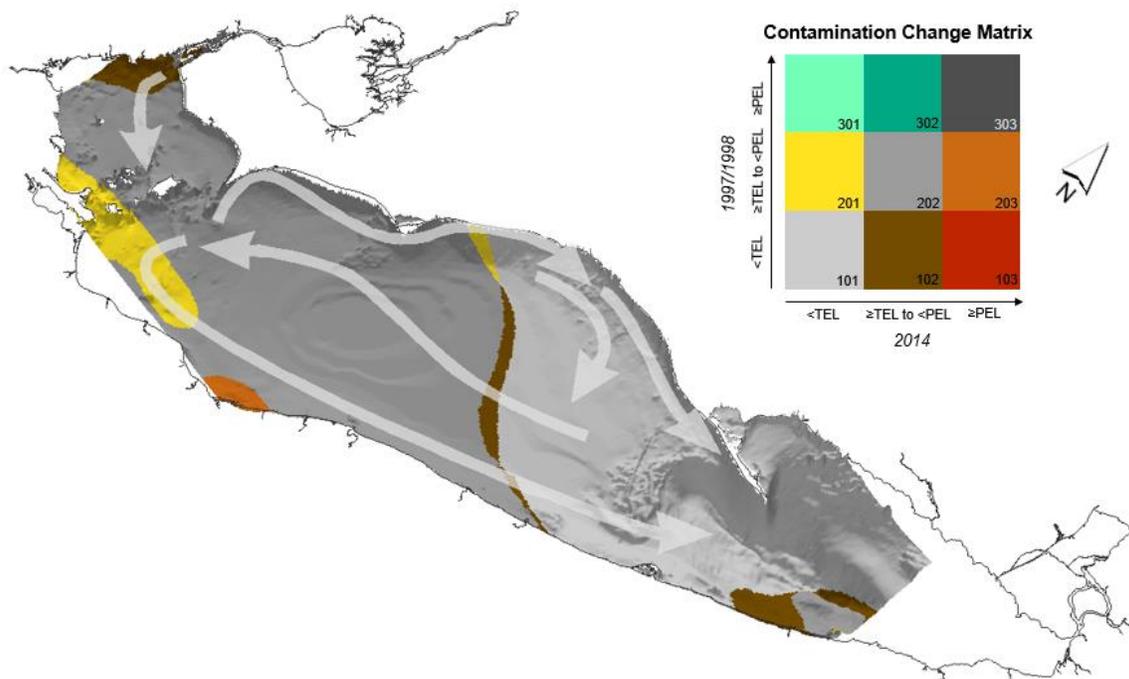


Figure 4-11: Change in Pb (log-normal) sediment contamination between 1997/1998 and 2014.

Table 4-9: Surface area change in Pb (log-normal) sediment contamination between 1997/1998 and 2014.

Change Category	Description	Area (km ²)	Area (%)
101	No Change, $<$ TEL	8,913.75	38.44
102	$<$ TEL to \geq TEL to $<$ PEL	1,088.75	4.70
201	\geq TEL to $<$ PEL to $<$ TEL	1,023	4.41
202	No change, \geq TEL to $<$ PEL	11,979.25	51.66
203	\geq TEL to $<$ PEL to \geq PEL	182	0.78
No Data		2,613.25	11.27

Change in Pb sediment contamination between historic (1971) and contemporary (2014) surveys detail a complex history of pollution in Lake Erie (Figure 4-12). Approximately 15,466 km² of the lake study area experienced change in Pb sediment contamination values between 1971 and 2014 (Table 4-10); 97.75% of which represented a decline in Pb contamination levels.

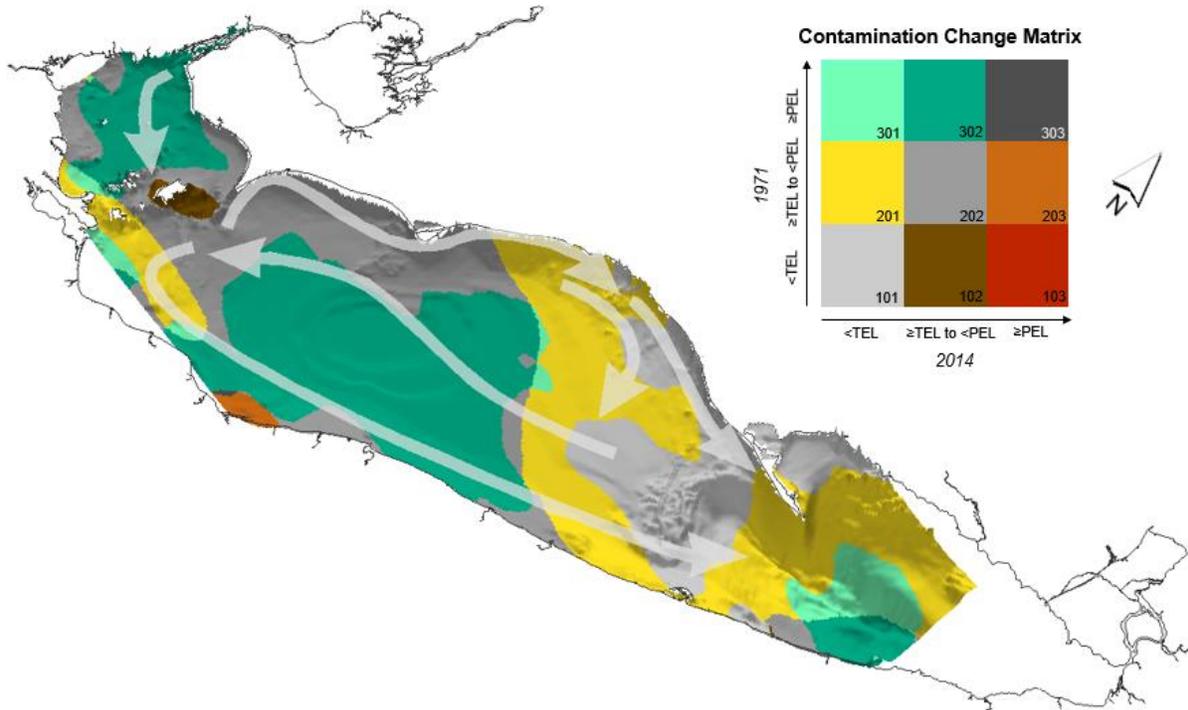


Figure 4-12: Change in Pb (log-normal) sediment contamination between 2017 and 2014.

Table 4-10: Surface area change in Pb (log-normal) sediment contamination between 1971 and 2014.

Change Category	Description	Area (km ²)	Area (%)
101	No Change, <TEL	3,048.25	13.13
102	<TEL to ≥TEL to <PEL	179.5	0.77
201	≥TEL to < PEL to <TEL	5,964	25.69
202	No change, ≥TEL to < PEL	4,692.5	20.21
203	≥TEL to < PEL to ≥ PEL	167.5	0.72
301	≥ PEL to <TEL	925	3.98
302	≥ PEL to ≥TEL to <PEL	8,230	35.45
303	No Change, ≥ PEL	12.25	0.05
No Data		2,581	11.12

Pb contamination ≥PEL in 1971 was widely reduced by 2014 where 8,230 km² of the study area declined to measures between ≥TEL to <PEL, and 925 km² were dramatically lowered to <TEL.

Sediments ranging in contamination \geq TEL to $<$ PEL were also largely reduced to $<$ TEL values throughout all three basins. Approximately 33% of the lake study area did not experience change in Pb contamination levels between 1971 and 2014. A very small area (12.25 km²) outside of Cleveland measured contaminated sediments \geq TEL in both 1971 and 2014. Contamination \geq TEL to $<$ PEL remained largely in the western and central basins, whereas $<$ TEL contamination persisted in the central and eastern basins between the two surveys.

Discrete areas of sediment in Lake Erie became more polluted between 1971 and 2014. Outside of Cleveland, an area of 167.5 km² increased in Pb sediment contamination from \geq TEL to $<$ PEL in 1971 to \geq PEL by 2014. Around Pelee Island and the Pelee-Lorraine ridge, sediments increased in Pb contamination from $<$ TEL in 1971 to \geq TEL to $<$ PEL (representing 179.5 km² of the study area) by 2014. Increasingly contaminated sediments only represented an area of 347 km² or 1.5% of the lake study area between 1971 and 2014.

4.5 Discussion

4.5.1 Historic and Contemporary Pb Sediment Contamination

Assessing interpolated surfaces of historic and contemporary sediment surveys identified fluctuating conditions of Pb sediment contamination throughout Lake Erie over both space and time. An overall trend of decreasing contamination at the \geq PEL threshold occurred over the survey periods. Throughout the entire lake, contaminated sediments \geq PEL had decreased in surface area from 1971 (36.72%) to 1997/1998 (0%) to 2014 (0.71%). By 2014, sediment contamination \geq PEL had changed from a lake-wide problem to a localized condition outside of Cleveland. The relatively unchanging state of Pb sediment contamination from \geq TEL to $<$ PEL is concerning, considering its high proportional representation of lake surface area almost exclusively in the western portion of the lake near highly urbanized and industrialized cities. Sediments contaminated to Pb levels $<$ TEL fluctuated between sediment surveys. In 1971, only 12.97% of the entire lake area was represented by sediments $<$ TEL; this extent more than tripled by 1997/1998, where contaminated sediments $<$ TEL represented 43.85% of the entire lake area (although predominantly located in the eastern half). By 2014, this contamination classification decreased in surface area representation to 38.62% while continually dominating the eastern half of Lake Erie. The dividing line between sediments \geq TEL to $<$ PEL and $<$ TEL through the middle of the central basin reflects a larger scale duality of land use, economies and population density between Canada and the United States (Murphy, 1962; RPA, 2008).

A complex dynamic of environmental policies has shaped navigational dredging and harbour development practices throughout Lake Erie since the 1970s. Prior to the enactment of the River and Harbour and Flood Control Act of 1970 (USACE, 2012) routine dredging of sediments from shipping routes by the US Army Corps of Engineers (USACE) resulted in open water dumping of dredged materials into Lake Erie. Since 1994, the State of Michigan has banned open water disposal of dredged materials, including contaminated sediments (USACE, 2012); 55% of dredged materials are placed in near shore or upland disposal facilities, 44% is housed in confined disposal facilities (CDF) and 1% of uncontaminated dredged sediments are disposed of in open waters (USACE, 2012). From the mouth of the St. Clair River to its border with Ohio, the State of Michigan operates 5 CFD in Lake Erie (USACE, 2012). That being said, the Detroit River continues to be the primary source of non-point source pollution to the lake (Marvin et al., 2002).

Approximately 1.5 million tons of sediment is dredged annually from the Ohio shoreline and harbours (OEPA, 2016). In stark contrast to Michigan, the State of Ohio dumps 44% of all dredged materials into the open waters of Lake Erie (USACE, 2012). Several offshore disposal locations including Toledo, Sandusky, and Fairport (approximately 45 km east of Cleveland) were identified to contain Pb contaminated sediments \geq PEL from the 1971, and 2014 sediment surveys. Open lake disposal of dredged harbour sediments will be prohibited in the State of Ohio after July 2020 (OEPA, 2016); until then, contaminated sediments along the southern shoreline of Lake Erie will be continuously disturbed, increasing the rate and quantity of sediment resuspension and transportation throughout the lake.

4.5.2 Multivariate Geovisualization

Patterns of polluted sediments and their transport fate were given additional geophysical content by overlaying contamination layers and average circulation directions on top of the bathymetric data. When viewed in two-dimensions (2D) (Figures 4-13 to 4-15), the Pelee-Lorraine ridge and Long Point-Erie sill are indistinguishable from other features throughout the lake basin. From the 1971 3D geovisualization, sediment transport from the western to central basin appears to be constrained by the Pelee-Lorraine ridge. The 1997/1998 and 2014 3D geovisualizations show sediments being transported eastward around Pelee Island and its ridge towards the central basin. The Long Point-Erie sill presents another potential barrier to sediment movement. The strongest current along the southern shoreline of Lake Erie flows towards the eastern basin through a narrow channel south of the sill. The 1971, 1997/1998, and 2014 3D geovisualizations display low Pb contamination values along the elevated sill and increasing levels of contamination aligned with the average circulation direction along the southern shoreline near Erie, PA.

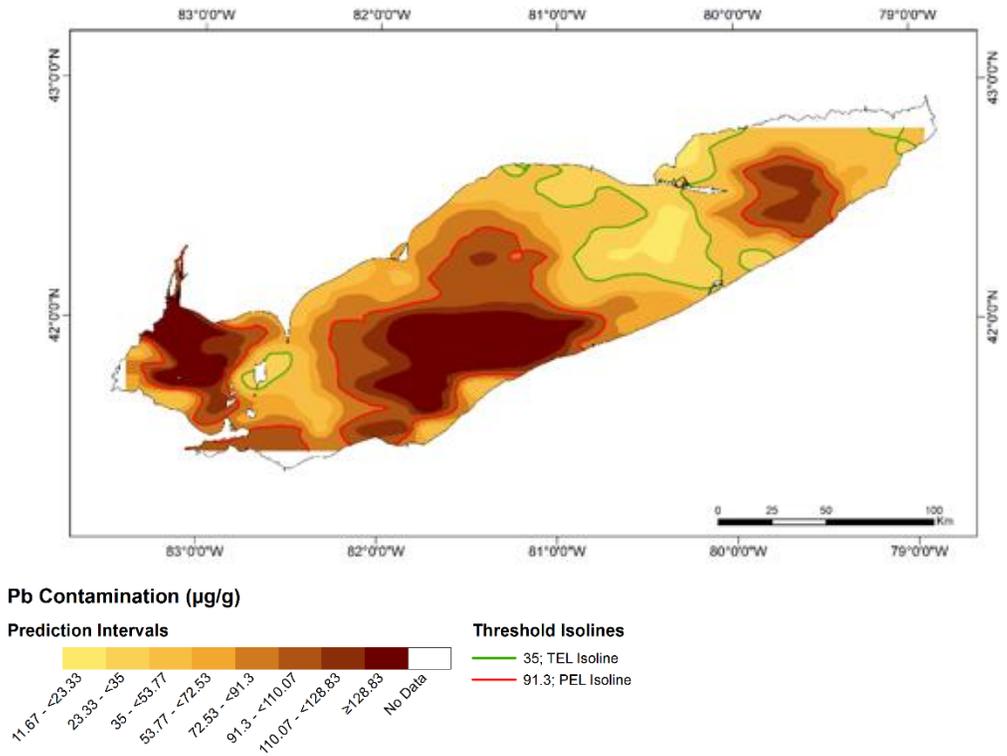


Figure 4-13: Kriging results of the 1971 Pb (log-normal) contamination in Lake Erie.

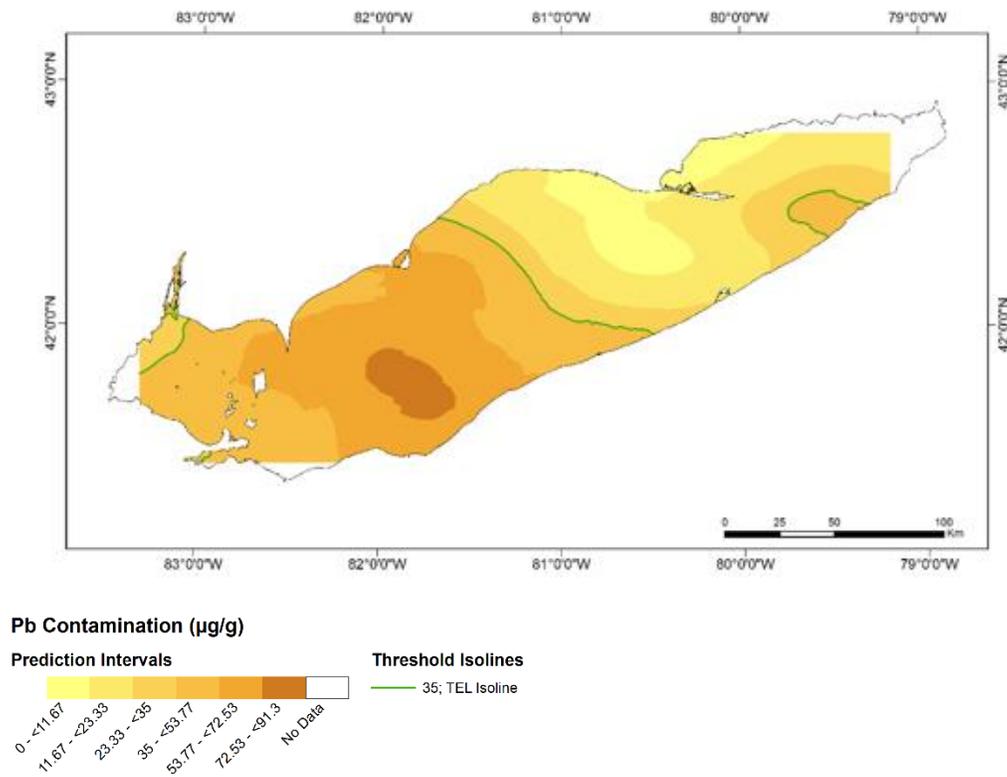


Figure 4-14: Kriging results of the 1997/1998 Pb (log-normal) contamination in Lake Erie.

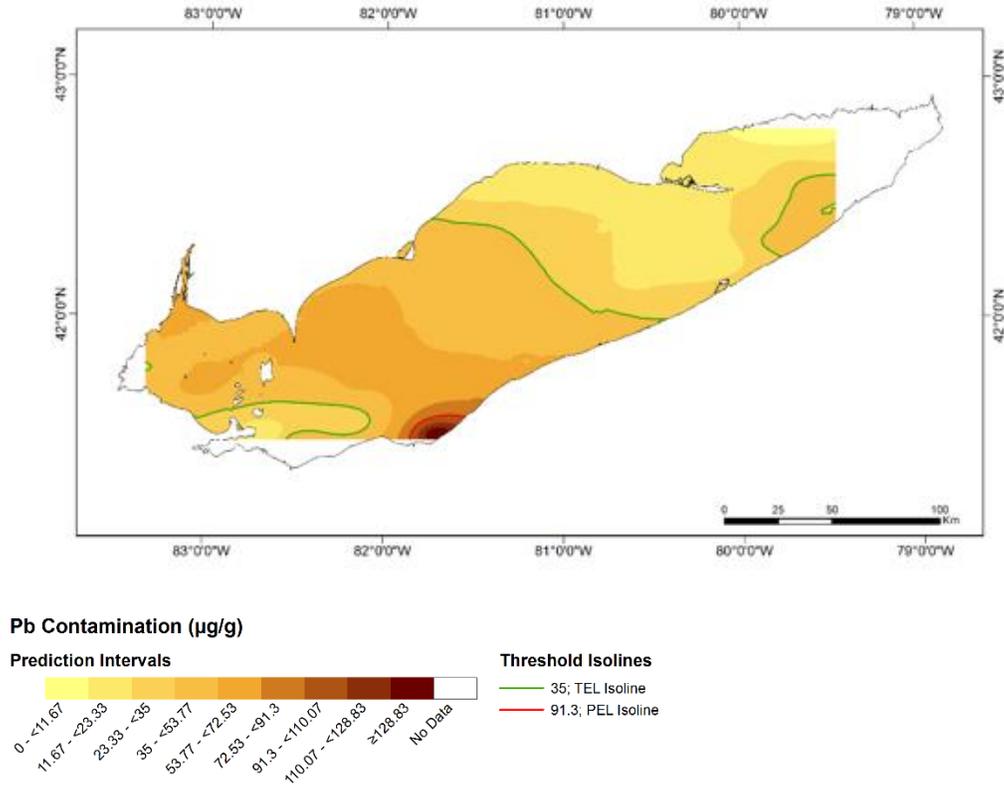


Figure 4-15: Kriging results of the 2014 Pb (log-normal) contamination in Lake Erie.

In 3D, the Pelee-Lorraine ridge appears as a barrier to sediment transport from the western to central basins in the 1971 geovisualization. In both the 1997/1998, and 2014 geovisualizations, sediments are transported eastward around Pelee Island and its ridge, along the Ontario shoreline near Point Pelee and into the central basin. The Long Point-Erie sill presents another potential barrier to sediment movement. The strongest lake current along the southern shoreline of Lake Erie flows towards the eastern basin through a narrow channel south of the sill. The 1971, 1997/1998, and 2014 3D geovisualizations display low Pb contamination values along the elevated sill; increasing levels of contamination aligned with the average circulation direction along the southern shoreline near Erie, Pennsylvania.

The steep sloping bathymetry into the middle of the eastern basin is also masked by 2D visualizations of the lake basin. When visualized in 3D, complex depth contours are revealed throughout the eastern basin with multiple levels of rapid changes in bathymetry. At the junction of the Long Point-Erie sill with the Long Point spit, and along the Ontario shoreline, very steep slopes (>89°) plunge to a maximum depth of 64 m. Along the New York state shoreline, the basins bathymetry is stepped with gradual declines into the depths of the eastern basin. Contaminated

sediments from all sampling surveys were predicted to be located on the gentler slopes of the eastern basin. Sediments along the steepest sides of the eastern basin are likely very unstable due to greater gravitational and erosional forces acting upon the steep slopes (Christopherson and Birkeland, 2015; Derosier et al., 2015).

4.6 Conclusion

Improvements to Lake Erie sediment contamination have been largely driven by improvements to environmental policies (USAEC, 2012). The overall decrease in highly contaminated sediments \geq PEL in 1997/1998 corresponds with the phasing out of leaded gasoline throughout the 1990s. Although sediments posing the greatest health risk to humans and aquatic life have been virtually eliminated by 2014, sediments \geq TEL to $<$ PEL dominate nearly half of the lake area and have the potential to pose environmental and health risks.

When viewed in 2D, contamination maps of Lake Erie neglect to inform the user of underlying geophysical characteristics of the lake basin. In 3D, bathymetric contours are illuminated from one perspective and shaded from another, encouraging the intuitive perception of changes in depth, slope and orientation based on changing colour contrast. The interpolated contamination maps from the 1971, 1997/1998, and 2014 sediment surveys are a valuable visual analytics tool for remediation efforts and contamination monitoring. Especially in the case of the Cleveland Harbour, remediation resources can be efficiently allocated to regions of the lake posing the greatest risk to the health and well-being of human and aquatic life. The water and sediments of Lake Erie know no political bounds; so long as shoreline regions operate to various environmental standards, the entire basin is at risk of sediment contamination.

Chapter 5: Conclusion

This dissertation contributes to the development of scientific communication methods for Great Lakes research and environmental decision-making in two ways. Analytical results first and foremost address the overall change in pollution patterns and quantities; as a proxy indicator for water quality, changes in sediment contamination patterns over time help to determine the overall health status of Lake Erie. The individual manuscripts (Chapters 2, 3, and 4) present novel research methodologies devised for improving the validity and overall usefulness of interpolated sediment contamination surfaces. Together, conclusions drawn from these chapters provide valuable research findings. Opportunities for practical applications of this research will be suggested along with recommendations for future research.

5.1 Status of the *Big Three* in Lake Erie

Sediment quality throughout Lake Erie has improved since the 1971 survey. Surface area coverage of contamination \geq PEL in 1971 (Hg: 50.16%; Cd: 13%; Pb: 37.92%) is nearly non-existent in predictions made by contemporary surveys (Hg 2014: 1.25%; Cd 1997/1998: 0%; Pb 2014: 0.78%). Overall poor sediment quality in 1971 (albeit, higher concentrations near industrial sources of contamination) have transformed into localized areas of concern, most notably outside of Cleveland, OH. In 1997/1998 Cd levels had reduced to \geq TEL to $<$ PEL, yet higher contamination levels ($<2.55 \mu\text{g/g}$) were found in lake sediments outside of Cleveland. In 2014, Hg and Pb contamination \geq PEL (≥ 0.697 and ≥ 128.83 , respectively) was located exclusively offshore of the Cleveland area. These regions of high Hg, Cd, and Pb sediment contamination are situated within Lake Erie's strongest wind-driven current. Heavily contaminated sediments will likely be transported eastward along the Ohio shoreline from prevailing currents and west/southwesterly winds (Thomas et al., 1976; Dusini et al., 2009).

The objective of the GLWQA is "to restore and maintain the chemical, physical, and biological integrity of the Great Lakes Basin Ecosystem" (IJC, 2002). In 2016, Adams and Stadler-Salt (of the EPA and ECCC, respectively) announced the status of sediment quality (as a sub-indicator of Lake Erie health) to be improving overall, aligning with spatiotemporal trends identified in Chapters 2, 3 and 4. Improvement in sediment quality may indirectly reflect regulatory initiatives (i.e. phase-out of leaded gasoline in Canada and the United States; phase-out of coal power plants in Ontario), and enhanced environmental stewardship of the 21st century (i.e. battery recycling programs). That being said, members of a complex binational, transboundary water quality agreement may be slow to react to necessary environmental changes. Contentious

dredging and open lake disposal practices will continue along the Ohio shoreline until 2020 (OEPA, 2016); a modern-day area of concern for all three non-essential heavy metals analyzed throughout this dissertation.

5.2 Validity of Spatially Interpolated Contamination Maps

Each of the three studies were built on a foundation of interpolated contamination surfaces. Aside from the inherent statistical validity of kriging interpolation, Chapters 2, 3, and 4 explored methods to improve the effectiveness of predicted sediment contamination maps for environmental decision making. The three inter-related studies share a common objective, to enhance the legitimacy and delivery of scientific communication methods. Driving this goal is the diverse group of stakeholders (government agencies, academic researchers, and public interest groups) working towards pollution control and environmental remediation of the Great Lakes.

In Chapter 2 the number of sampling points required to produce statistically valid and spatially accurate predicted data surfaces was questioned; and the answer was not as straightforward as initially anticipated. The density and pattern of sampling locations is just as important as the number of sampling points used in a kriging analysis. The research findings made in Chapter 2 owe to the larger discussion of map validity in this dissertation. Where the 1971 Hg survey was considered to be the best dataset from which to interpolate sediment contamination, several randomized down-sampled subsets also produced statistically valid and meaningful pollution maps. Subsets from 90% to 20% of the complete 1971 Hg dataset had high enough sampling densities to maintain some grid-like pattern. Interpolated surfaces produced from these subsets predicted similar contamination patterns (although increasingly generalized) to that of the complete dataset. Sampling densities ranging from 0.01 to 0.002 locations/km² are enough to produce statistically valid and meaningful sediment contamination maps. Therefore, consistency in the design and sampling pattern of a sediment survey is just as important as the number of locations sampled. With limited resources, contemporary surveys aim to sample locations of non-point source pollution, which are not likely to follow the grid-like pattern of historic surveys. By limiting the spatial extent of the kriging interpolation, or integrating high-density auxiliary variables employed in Chapter 3 and 4, smaller sediment surveys may produce valid sediment contamination maps on a budget.

Improving the validity of interpolated surfaces with auxiliary variables and alternative classification indices was examined in Chapter 3. Following the fundamentals of co-kriging, integrating densely sampled sediment properties with low density survey samples successfully enhanced the spatial

understanding of contamination patterns over time. In conjunction with the Index of Geoaccumulation, anthropogenically sourced Cd was differentiated from natural occurrences of the heavy metal. Naturally occurring Cd ($<1 \mu\text{g/g}$) was identified largely in coarse grained glacial till while fine-grained clay and silt were dominated by anthropogenic Cd contamination in both the 1971 and 1997/1998 surveys. Here, limited survey data have been successfully enhanced by supplementing lake-wide sediment substrate information. Similarly, to Chapter 2, the results of this study foster more informed decision making on spatial aspects of pollution control efforts.

Bathymetric data used in Chapter 4 are arguably the most innovative auxiliary variable included in the multi-variable analysis of sediment contamination in Lake Erie. Used mainly as a visualization tool, bathymetric data also supplemented the low sampling density of the 1997/1998 and 2014 sediment surveys with depth information. Pb contamination patterns display similar temporal trends to Hg and Cd contamination examined in Chapters 2 and 3. When viewed in 2D, the relationship between sediment contamination patterns, hydrodynamic processes and geophysical features in the lake are not clearly visible. Visualizing contamination patterns in 3D, along with average circulation, enhances the understanding of contamination transport throughout the lake basin. Similar to research findings made in Chapter 3, multi-variate analyses of sediment contamination with circulation patterns and depth better illustrate pollution patterns relative to bathymetric contours of the lake. Weaker currents along the less developed Ontario shoreline display consistently low levels of Pb contamination in 1971, 1997/1998, and 2014. Moderate ($\geq\text{TEL}$ to $<\text{PEL}$) to high ($\geq\text{PEL}$) levels of Pb contamination in the eastern basin have likely been transported from industrial sources of pollution along the Ohio shoreline by stronger currents. By visualizing sediment contamination patterns in 3D, it is easier to assess the likely transportation pathways of heavily polluted sediments throughout the Lake Erie basin. Similar to the research methodologies devised in Chapters 2 and 3, this innovative geospatial technique creates significant opportunities for valuable scientific communication amongst various stakeholders.

5.3 Avenues of Future Research Applications

This growing body of research is a valuable resource to assist in assessing the performance of multi-stakeholder plans, policies and programs towards mitigating or preventing further heavy metal pollution into the Great Lakes Basin. Identifying the change in contamination quantities and patterns throughout Lake Erie can support discussions of progress under the Lake Erie Lakewide Action and Management Plans (LaMPs), Remedial Action Plans (RAPs) for Areas of Concern (AOC) and the development of jurisdictional regulations.

Maps and 3D geovisualizations are excellent tools for communicating spatial trends to a wide variety of individuals, ranging from the general public to scientists and government officials. Sediment samples are a snapshot of contamination levels in a specific location, at a certain time. In point form, the value of sediment contamination data to pollution control efforts and scientific communication is hindered by the disjunctive, and incomplete nature of traditional dot and/or proportional circle maps. Spatial interpolation methods, like kriging, are the first step to enhancing the communicative value of contamination information from point data to a continuous data layer.

Conceptually, the research methodologies created through empirical geospatial analyses in Chapters 2, 3 and 4 may help to bridge the knowledge gap between Great Lake experts and Great Lake citizens. Every three years, the International Joint Commission (IJC) reports on the Status of Great Lakes Ecosystem Integrity, holds a public forum, and reports on the progress towards binational commitments under the GLWQA (IJC, 2016). Maps and research available to the general public allow individuals to educate themselves on the evolving status of Great Lakes sediment contamination, increasing informed public participation at triennial IJC public forums. Research methodologies integrating ancillary variables may be operationalized with any sediment contaminant, so long as sediment properties and bathymetric data are available for the lake under analysis. Moving forward, it is essential to create maps of as many persistent toxic chemicals and heavy metals as possible for each lake to develop a thorough and historically significant body of knowledge for mitigating Great Lakes sediment contamination into the future.

References

- 3D Analyst: 3D Analyst and ArcScene. 2017. ArcGIS Desktop. Available online: www.desktop.arcgis.com/en/arcmap/latest/extensions/3d-analyst/3d-analyst-and-arcscene.htm. [Accessed: July 6, 2017].
- Adams, J., and N. Stadler-Salt. How are the Great Lakes Doing? Presented at the Great Lakes Public Forum. October 4-6, 2016. Toronto, Ontario, Canada. Available online: www.binational.net/wp-content/uploads/2016/11/01%20SOGL%20Presentation%20ENG.pdf. [Accessed: August 22, 2017].
- Adriaens, P., S. Batterman, J. Blum, K. Hayes, P. Meyers, and W. Weber. 2002. Great Lakes Sediment: Contamination, Toxicity and Beneficial Re-Use. White Paper Commissioned by Michigan Sea Grant and the School for Natural Resources and the Environment (SNRE), Chicago State University: Illinois, USA. Available online: www.csu.edu/cerc/researchreports/documents/GreatLakesSedimentsContaminationToxicityBeneficialReUse.pdf. [Accessed: July 16, 2017].
- Agency for Toxic Substances and Disease Registry (ATSDR). 2007. Toxicological profile for Lead. Available online: www.atsdr.cdc.gov/ToxProfiles/tp.asp?id=96&tid=22. [Accessed: June 24, 2017].
- Alther, G.R. 1981. Transport of Dredged Sediments After Disposal Operational in Lake Erie. *The Ohio Journal of Science*. Vol. 81, No. 1, pp. 2-8.
- Andrienko, N. and G. Andrienko. 2003. Informed Spatial Decisions through Coordinated Views. *Information Visualization*. Vol. 2, No. 4, pp. 270-285.
- ArcGIS for Desktop. 2016. Performing Cross-Validation and Validation. ArcMap 10.4. Available online: www.desktop.arcgis.com/en/arcmap/latest/extensions/geostatistical-analyst/performing-cross-validation-and-validation.htm. [Accessed: July 21, 2016].
- Ashworth, W. 1987. The Late, Great Lakes: An Environmental History. Wayne State University Press: Michigan, USA. 288pp.
- Bai, X., J. Wang, D.J. Schwab, Y. Yang, L. Luo, G.A. Leshkevich, and S. Lui. 2013. Modeling 1993-2008 Climatology of Seasonal General Circulation and Thermal Structure in the Great Lakes using FVCOM. *Ocean Modelling*. Vol. 65, pp. 40-63.
- Beletsky, D., J.H. Saylor, and D.J. Schwab. 1999. Mean Circulation in the Great Lakes. *Journal of Great Lakes Research*. Vol. 25, No. 1, pp. 78-93.
- Birch, G.F., T.J. Gunns, and M. Olmos. 2015. Sediment-bound Metals as Indicators of Anthropogenic Change in Estuarine Environments. *Marine Pollution Bulletin*. Vol. 101, No. 1, pp. 243-257.
- Blais, J.M., and J. Kalff. 1995. The Influence of Lake Morphometry on Sediment Focusing. *Limnology and Oceanography*. Vol. 40, No. 3, pp. 582-588.

- Bogue, M.B. 2000. *Fishing the Great Lakes: An Environmental History, 1783-1933*. The University of Wisconsin Press: Wisconsin, USA. 464 pp.
- Burrough, P.A. and R.A. McDonnell. 1998. *Principles of Geographical Information Systems*. Oxford University Press: Oxford. 333pp.
- Canadian Council of Ministers of the Environment (CCME). 1999. *Canadian Environmental Quality Guidelines*. Winnipeg, MB: Canadian Council of Ministers of the Environment.
- Canadian Council of Ministers of the Environment (CCME). 2003. *Climate, Nature, People: Indicators of Canada's Changing Climate*. Climate Change Indicators Task Group of the Canadian Council of Ministers of the Environment. Winnipeg, MB. 51 pp.
- Centers for Disease Control and Prevention (CDC). 2014. Lead: Prevention Tips. Available online: www.cdc.gov/nceh/lead/tips.htm. [Accessed: July 16, 2017].
- Chang, Y.H., M.D. Scrimshaw, R.H.C. Emmerson, and J.N. Lester. 1998. Geostatistical Analysis of Sampling Uncertainty at the Tollesbury Managed Retreat Site in Blackwater Estuary, Essex, UK: Kriging and Cokriging Approach to Minimise Sampling Density. *Science of the Total Environment*. Vol. 221, No. 1, pp. 43-57.
- Chen, G., K. Zhao, G.J. McDermid, and G.J. Hay. 2012. The Influence of Sampling Density on Geographically Weighted Regression: a Case Study using Forest Canopy Height and Optical Data. *International Journal of Remote Sensing*. Vol. 33, No. 9, pp. 2909-2924.
- Christopherson, R.W., and G.H. Birkeland. 2015. *Geosystems: An Introduction to Physical Geography (9th Ed.)*. New York, NY: Pearson Education Inc. 688 pp.
- Clark, I. 1982. *Practical Geostatistics*. Applied Science Publishers Ltd.: Essex, England. 129pp.
- Dalia, M.S., A. Salem, A. Khaled, A. El Nemr, and A. El-Sikaily. 2014. Comprehensive Risk Assessment of Heavy Metals in Surface Sediments along the Egyptian Red Sea Coast. *The Egyptian Journal of Aquatic Research*. Vol. 40, No. 4, pp. 349-362.
- Dartmouth Toxic Metals: Superfund Research Program (Dartmouth). 2012. A Metals Primer. Dartmouth College. Available online: www.dartmouth.edu/~toxmetal/toxic-metals/more-metals/cadmium-faq.html. [Accessed: July 16, 2016].
- Derosier, A.L., S.K. Hanshue, K.E. Wehrly, J.K. Farkas, and M.J. Nichols. 2015. Michigan's Wildlife Action Plan. Michigan Department of Natural Resources. Lansing, MI. Available online: www.michigan.gov/documents/dnr/05_wap_littoral_zones_500065_7.pdf. [Accessed: July 31, 2017].
- De Solla, S.R., J. Struger, and T.V. McDaniel. 2012. Detection Limits can Influence the Interpretation of Pesticide Monitoring Data in Canadian Surface Waters. *Chemosphere*. Vol. 86, No. 6, pp. 565-571.
- Dille, J.A., M. Milner, J.J. Groeteke, D.A. Mortensen, and M.M. Williams II. 2002. How Good is your Weed Map? A Comparison of Spatial Interpolators. *Weed Science*. Vol. 51, No. 1, pp. 44-55.

- Duarte, C.M., and J. Kalff. 1986. Littoral Slope as a Predictor of the Maximum Biomass of Submerged Macrophyte Communities. *Limnology and Oceanography*. Vol. 31, No. 5, pp. 1072-1080.
- Dusini, D.S., D.L. Foster, J.A. Shore, and C. Merry. 2009. The Effect of Lake Erie Water Level Variations on Sediment Resuspension. *Journal of Great Lakes Research*. Vol. 35, No. 1, pp. 1-12.
- Esri. 2017. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.
- Essex Region Source Protection Area (ERSPA). 2015. State of Climate Change Research in the Great Lakes Region. Updated Assessment Report – Essex Region Source Protection Area. Available online: www.essexregionsourcewater.org/downloads/2015/chapter-6-state-of-climate-change-research-in-the-great-lakes-region.pdf. [Accessed: July 16, 2017].
- Evers, D.C., J.G. Wiener, C.T. Driscoll, D.A. Gay, N. Basu, B.A. Monson, K.F. Lambert, H.A. Morrison, J.T. Morgan, K.A. Williams, A.G. Soehl. 2011. Great Lakes Mercury Connections: The Extent and Effects of Mercury Pollution in the Great Lakes Region. Biodiversity Research Institute: Gorham, Maine. Report BRI 2011-18. pp. 44.
- Forsythe, K.W., M. Dennis, and C.H. Marvin. 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 Research Journal of Canada*. Vol. 39, No. 3, pp. 190-206.
- Forsythe, K.W., A. Gawedzki, P.S. Rodriguez, K.N. Irvine, and M.F. Perrelli. 2013. Geospatial Estimation of Mercury Contamination in Buffalo River Sediments. *Soil and Sediment Contamination An International Journal*. Vol. 22, No. 5, pp. 521-531.
- Forsythe, K.W., K.N. Irvine, D.M. Atkinson, M. Perelli, J.M. Aversa, S.J. Swales, A. Gawedzki, and D.J. Jakubek. 2015. Assessing Lead Contamination in Buffalo River Sediments. *Journal of Environmental Informatics*. Vol. 26, No. 2, pp. 106-111.
- Forsythe, K.W. and C.H. Marvin. 2005. Analyzing the Spatial Distribution of Sediment Contamination in the Lower Great Lakes. *Water Quality Research Journal of Canada*. Vol. 40, No. 4, pp. 389-401.
- Forsythe, K.W. and C.H. Marvin. 2009. Assessing Historical Versus Contemporary Mercury and Lead Contamination in Lake Huron Sediments. *Aquatic Ecosystem Health and Management*. Vol. 12, Issue 1, pp. 101-109.
- Forsythe, K.W., C.H. Marvin, D.E. Mitchell, J.M. Aversa, S.J. Swales, D.A. Burniston, J.P. Watt, D.J. Jakubek, M.H. McHenry, R.R. Shaker. 2016c. Utilization of Bathymetry Data to Examine Lead Sediment Contamination Distributions in Lake Ontario. *AIMS Environmental Science*. Vol. 3, No. 3, pp. 347-361.
- Forsythe, K.W., C.H. Marvin, C.J. Valancius, J.P. Watt, J.M. Aversa, S.J. Swales, D.J. Jakubek, and R.R. Shaker. 2016b. Geovisualization of Mercury Contamination in Lake St. Clair Sediments. *Journal of Marine Science and Engineering*. Vol. 4, No. 19, pp. 1-10.

- Forsythe, K.W., C.H. Marvin, C.J. Valancius, J.P. Watt, S.J. Swales, J.M. Aversa, D.J. Jakubek. 2016a. Using Geovisualization to Assess Lead Sediment Contamination in Lake St. Clair. *The Canadian Geographer*. Vol. 60, No. 1, pp. 149-158.
- Forsythe, K.W., K. Paudel, and C.H. Marvin. 2010. Geospatial Analysis of Zinc Contamination in Lake Ontario Sediments. *Journal of Environmental Informatics*. Vol. 16, No. 1, pp. 1-10.
- Fujisaki, A., J. Wang, H. Hu, D.J. Schwab, N. Hawley, Y.R. Rao. 2012. A Modeling Study of Ice-Water Processes for Lake Erie Applying Coupled Ice-Circulation Models. *Journal of Great Lakes Research*. Vol. 38, No. 4, pp. 585 – 599.
- Gawedzki, A. and K.W. Forsythe. 2012. Assessing Anthracene and Arsenic Contamination within Buffalo River Sediments. *International Journal of Ecology*. Vol. 2012, pp. 1-9.
- Goovaerts, P. 1999. Geostatistics in Soil Science: State-of-the-art and Perspectives. *Geoderma*. Vol. 89, No. 1-2, pp. 1-45.
- Gotway, C.A., R.B. Ferguson, G.W. Hergert and T.A. Peterson. 1996. Comparison of Kriging and Inverse-Distance Method for Mapping Soil Parameters. *Soil Science Society of America Journal*. Vol. 60, No. 4, pp. 1237-1247.
- Gotway, C. A., and A. H. Hartford. 1996. Geostatistical Methods for Incorporating Auxiliary Information in the Prediction of Spatial Variables. *Journal of Agricultural, Biological, and Environmental Statistics*. Vol. 1, No. 1, pp. 17–39.
- Great Lakes Information Network (GLIN). 2004. Great Lakes-St. Lawrence water flows. Available online: www.great-lakes.net/envt/water/levels/flows.html. [Accessed: July 14, 2016].
- Great Lakes Information Network (GLIN). 2016. Lake Erie: Facts and Figures. Available online: www.great-lakes.net/lakes/ref/eriefact.html. [Accessed: June 24, 2017].
- Great Lakes Information Network (GLIN). 2017. Great Lakes Facts and Figures. Available online: www.great-lakes.net/lakes/ref/lakefact.html. [Accessed: August 13, 2017].
- Gupta, S., V. Jena, N. Matic, V. Kapralova, and J.S. Solanki. 2014. Assessment of Geo-accumulation Index of Heavy Metal and Source of Contamination by Multivariate Factor Analysis. *International Journal of Hazardous Materials*. Vol. 2, No. 2, pp. 18-20.
- Habashi, F. 2011. Pollution Problems in the Metallurgical Industry: A Review. *Journal of Mining and Environment*. Vol. 2, No. 1, pp. 17-26.
- Haltuch, M.A., and P.A. Berkman. 2000. Geographic Information System (GIS) Analysis of Ecosystem Invasion: Exotic Mussels in Lake Erie. *Limnology and Oceanography*. Vol. 45, No. 8, pp. 1778-1787.
- Hawley, N. and B.J. Eadie. 2007. Observations of Sediment Transport in Lake Erie during the Winter of 2004-2005. *Journal of Great Lakes Research*. Vol. 33, No. 4, pp. 816-827.

- Health Canada. 2009. Lead Information Package – Some Commonly Asked Questions About Lead and Human Health. Available online: www.canada.ca/en/health-canada/services/environmental-workplace-health/environmental-contaminants/lead/lead-information-package-some-commonly-asked-questions-about-lead-human-health.html. [Accessed: July 16, 2017].
- Hessl, A., J. Miller, J. Kernan, D. Keenum, D. McKenzie. 2007. Mapping Paleo-fire Boundaries from Binary Point Data: Comparing Interpolation Methods. *The Professional Geographer*. Vol. 59, No. 1, pp. 87-104.
- IBM Corporation. 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corporation.
- Idriss, A.A., and A.K. Ahmad. 2012. Heavy Metal Concentrations (Cu, Cd, and Pb) in sediments in the Juru River, Penang, Malaysia. *Journal of Biological Sciences*. Vol. 12, No. 7, pp. 376-384.
- International Joint Commission (IJC). 1972. Great Lakes Water Quality Agreement, with Annexes and Texts and Terms of References, Between the United States of America and Canada. Available online: www.ijc.org/files/publications/C23.pdf. [Accessed: July 16, 2017].
- International Joint Commission (IJC). 1982. First Biennial Report under the Great Lakes Water Quality Agreement of 1978. Available online: www.ijc.org/files/publications/ID604.pdf. [Accessed: July 16, 2017].
- International Joint Commission (IJC). 1994. Seventh Biennial Report under the Great Lakes Water Quality Agreement of 1978 to the Governments of the United States and Canada and the States and Provinces of the Great Lakes Basin. Available online: www.ijc.org/files/publications/seventh-biennial-report-under-glwqa-ijc.pdf. [Accessed: July 16, 2017].
- International Joint Commission (IJC). 2002. Eleventh Biennial Report on Great Lakes Water Quality: The Challenge to Restore and Protect the Largest Body of Fresh Water in the World. Available online: www.ijc.org/files/publications/C89.pdf. [July 16, 2017].
- International Joint Commission (IJC). 2016. Great Lakes Water Quality Agreement. Protecting Shared Resources. Available online: www.ijc.org/en_/Great_Lakes_Quality. [Accessed: August 25, 2017].
- Jain, C.K. and D. Ram. 1997. Adsorption of metal ions on bed sediments. *Hydrological Sciences Journal*. Vol. 42, No. 5, pp. 713-723.
- Jakubek, D.J., and K.W. Forsythe. 2004. A GIS-based Kriging Approach for Assessing Lake Ontario Sediment Contamination. *The Great Lakes Geographer*. Volume 11, No. 1, 2004, pp. 1-14.
- Johnston K., J.M. Ver Hoef, K. Krivoruchko, and N. Lucas. 2003. Using ArcGIS geostatistical analyst. ESRI, Redlands, USA. Available online: www.dusk2.geo.orst.edu/gis/geostat_analyst.pdf. [Accessed: February 9, 2017].

- Kabata-Pendias, A., and A.B. Mukherjee. 2007. Trace Elements from Soil to Human. Springer Science and Business Media. 550 pp.
- Kersten, M. and F. Smedes. 2002. Normalization Procedures for Sediment Contaminants in Spatial and Temporal Trend Monitoring. *Journal of Environmental Monitoring*. Vol. 4, No. 1, pp. 109-115.
- Khosh Eghbal, M.Z. 2014. Spatial Distribution of Sediment Pollution in the Khajeh Kory River using Kriging and GIS. *Earth Sciences Research Journal*. Vol. 18, No. 2, pp. 173-179.
- Knust, C., and M.F. Buchroithner. 2014. Principles and Terminology of True-3D Geovisualization. *The Cartographic Journal*. Vol. 51, No. 3, pp. 191-202.
- Kot-Wasik, A. 2005. Water Quality Control: Part III, Human Impact on Water Resources. Politechnika Gdanska, Poland. Available online: www.pg.gda.pl/chem/Dydaktyka/Analityczna/WQC/wqc_p3.pdf. [Accessed: July 16, 2016].
- Koua, E.L., A. Maceachren, and M-J. Kraak. 2006. Evaluating the Usability of Visualization Methods in An Exploratory Geovisualization Environment. *International Journal of Geographical Information Science*. Vol. 20, No. 4, pp. 425-448.
- Krivoruchko, K. (2005). Introduction to Modeling Spatial Process using Geostatistical Analyst. Redlands: Environmental Systems Research Institute. 27pp.
- Kumar, C., W. Heuten, S. Boll. 2013. Geovisualization for End User Decision Support: Easy and Effective Exploration of Urban Areas. GeoViz_Hamburg 2013: Interactive Maps That Help People Think. Available online: www.geomatik-hamburg.de/geoviz/online/Kumar_Abstract.pdf. [Accessed: June 29, 2017].
- Largueche, F-Z.B. 2006. Estimating Soil Contamination with Kriging Interpolation Method. *American Journal of Applied Sciences*. Vol. 3, No. 6, pp. 1894-1898.
- Lark, R.M. and R.B. Ferguson. 2004. Mapping risk of soil nutrient deficiency or excess by disjunctive indicator kriging. *Geoderma*. Vol. 118, No. 1-2, pp. 39-53.
- Lewis, M. 1966. Sedimentation Studies of Unconsolidated Deposits in the Lake Erie Basin. Ph.D. thesis. University of Toronto.
- Li, J., and A.D. Heap. 2008. A Review of Spatial Interpolation Methods for Environmental Scientists. Geoscience Australia: Canberra, Australia. 137pp.
- Lick, W., J. Lick, and C.K. Ziegler. 1994. The Resuspension and Transport of Fine-Grained Sediments in Lake Erie. *Journal of Great Lakes Research*. Vol. 20, No. 4, pp. 599-612.
- Lloyd, D. and J. Dykes. Human-Centred Approaches in Geovisualization Design: Investigating Multiple Methods Through a Long-Term Case Study. 2011. *IEEE Transactions on Visualization and Computer Graphics*. Vol. 17, No. 12, pp. 2498-2507.

- Lui, X., J. Wu, J. Xu. 2006. Characterizing the Risk Assessment of Heavy Metals and Sampling Uncertainty Analysis in Paddy Field by Geostatistics and GIS. *Environmental Pollution*. Vol. 141, No. 2, pp. 257-264.
- Luoma, S.N. 1989. Can we Determine the Biological Availability of Sediment-bound Trace Elements? *Hydrobiologia*. Vol. 176, No. 1, pp. 379-396.
- MacEachren, A.M., M. Gahegan, W. Pike, I. Brewer, G. Cai, E. Lengerich, and F. Hardisty. 2004. Geovisualization for Knowledge Construction and Decision Support. *IEEE Computer Graphics and Applications*. Vol. 24, No. 1, pp. 13-17.
- Markus, J. and A.B. McBratney. 2001. A Review of the Contamination of Soil with Lead 2. Spatial Distribution and Risk Assessment of Soil Lead. *Environment International*. Vol. 27, No. 5, pp. 399-411.
- Marvin C.H., M.N. Charlton, E.J. Reiner, T. Kolic, K. MacPherson, G.A. Stern, E. Braekevelt, J.F. Estenik, L. Thiessen, and S. Painter. 2002. Surficial Sediment Contamination in Lake Erie and Ontario: A Comparative Analysis. *Journal of Great Lakes Research*. Vol. 28, No. 3, pp. 437-450.
- Marvin C.H., M.N. Charlton, G.A. Stern, E. Braekevelt, E.J. Reiner, and S. Painter. 2003. Spatial and temporal trends in sediment contamination in Lake Ontario. *Journal of Great Lakes Research*. Vol. 29, No. 2, pp. 317-331.
- Marvin, C., L. Grapentine, and S. Painter. 2004a. Application of a Sediment Quality Index to the Lower Laurentian Great Lakes. *Environmental Monitoring and Assessment*. Vol. 91, No. 1-3, pp. 1-16.
- Marvin C.H., S. Painter, and R. Rossmann. 2004. Spatial and temporal patterns in mercury contamination in sediments of the Laurentian Great Lakes. *Environmental Research*. Vol. 95, No. 3, pp. 351-362.
- McBratney, A.B. and R. Webster. 1986. Choosing Functions for Semi-variograms of Soil Properties and Fitting them to Sampling Estimates. *Journal of Soil Science*. Vol. 37, No. 4, pp. 617-639.
- McCormick, B.H., T.A. DeFanti, and M.D. Brown. 1987. Visualization in Scientific Computing. *Computer Graphics*. Vol. 21, No. 6, pp. 1-14.
- McKenney, D.W., R.S. Rempel, L.A. Venier, Y. Wang, and A.R. Bisset. 1998. Development and Application of a Spatially Explicit Moose Population Model. *Canadian Journal of Zoology*. Vol., 76, No. 10, pp. 1922-1931.
- Miller, J.A. 1998. Confined Disposal Facilities on the Great Lakes. U.S. Army Corps of Engineers, Great Lakes & Ohio River Division. Available online: www.westernlakeerie.org/open_lake_corps_cdf_sixe_cost_rules_etc.pdf. [Accessed: June 4, 2017].

- Montgomery, D.C., E.A. Peck, and G.G. Vining. 2001. Introduction to Linear Regression Analysis. John Wiley and Sons. New York: New York. Available online: <http://www.math.nsysu.edu.tw/~lomn/homepage/class/92/DurbinWatsonTest.pdf>. [Accessed: October 21, 2017].
- Morang, A., M.C. Mohr, and C.M. Forgette. 2011. Longshore Sediment Movement and Supply along the U.S. of Lake Erie. *Journal of Coastal Research*. Vol. 27, No. 4, pp. 619-635.
- Müller, G. 1969. Index of Geo-accumulation in Sediments of the Rhine River. *Geo Journal*. Vol. 24, No. 2, pp. 108-118.
- Müller, G. 1981. Die Schwermetallbelastung der Sedimenten des Neckars und Seiner Nebenflüsse. *Chemiker-Zeitung*. No. 6, pp. 157-164.
- Murphy, R.E. 1962. Reviewed work: Megalopolis: The Urbanized Northeastern Seaboard of the United States by Jean Gottmann. *American Geographical Society, Geographical Review*. Vol. 52, No. 3, pp. 452-454.
- National Geophysical Data Center (NGDC). 1999. Bathymetry of Lake Erie and Lake Saint Clair. National Geophysical Data Center, NOAA. doi:10.7289/V5KS6PHK.
- National Institute of Standards and Technology (NIST). 2013. e-Handbook of Statistical Methods. Engineering Statistics Handbook. U.S. Department of Commerce. Available online: www.itl.nist.gov/div898/handbook/. [Accessed: July 16, 2017].
- National Oceanic and Atmospheric Administration (NOAA). 2014. Great Lakes Bathymetry. Available online: www.ngdc.noaa.gov/mgg/greatlakes/greatlakes.html. [Accessed: June 26, 2017].
- O'Grady, K., and A. Perron. 2011. Reformulating Lead-Based Paint as a Problem in Canada. *American Journal of Public Health*. Vol. 101, No. 1, pp. 176 – 187.
- Ohio Environmental Protection Agency. n.d. Lake Erie Dredged Material Program. Dredged Material: Improving Ohio's Water Quality and Economy. Available online: www.epa.ohio.gov/dir/dredge.aspx. [Accessed: July 31, 2017].
- Olea, R.A. 1984. Sampling Design Optimization for Spatial Functions. *Mathematical Geology*. Vol. 16, No. 4, pp. 369-392.
- Osburn, W.L. 2000. Geostatistical analysis: potentiometric network for the Upper Floridian Aquifer in the St. Johns River water management district. Available online: www.sjrwmd.com/technicalreports/pdfs/TP/SJ2000-1.pdf [Accessed: May 11, 2017].
- Ouyang, Y., P. Nkedi-Kizza, R.S. Mansell, J.Y. Ren. 2003. Spatial Distribution of DDT in Sediments from Estuarine Rivers of Central Florida. *Journal of Environmental Quality*. Vol. 32, No. 5, pp. 1710-1716.
- Pachana, K., A. Wattanakornsiri, and J. Nanuam. 2010. Heavy Metal Transport and Fate in the Environmental Compartments. *NU Science Journal*. Vol. 7, No. 1, pp. 1-11.

- Painter, S., C. Marvin, F. Rosa, T.B. Reynoldson, M.N. Charlton, M. Fox, P.A. Lina Thiessen, and J.F. Estenik. 2001. Sediment Contamination in Lake Erie: A 25-Year Retrospective Analysis. *Journal of Great Lakes Research*. Vol. 27, No. 4, pp. 434-448.
- Pincus, H.J. 1953. The Motion of Sediment Along the South Shore of Lake Erie. *Coastal Engineering Proceedings*. No. 4, pp. 119-146.
- Pintilie, S., L. Branza, C. Betianu, L.V. Pavel, F. Ungureanu, M. Garvilescu. 2007. Modelling and Simulation of Heavy Metals Transport in Water and Sediments. *Environmental Engineering and Management Journal*. Vol. 6, No. 2, pp. 153-161.
- Prokop, Z., P. Cupr, V. Zlevorova-Zlamalikova, J. Komarek, L. Dusek, and I. Holoubek. 2003. Mobility, Bioavailability, and Toxic Effects of Cadmium in Soil Samples. *Environmental Research*. Vol. 91, No. 2, pp. 119-126.
- Ramsey, M.H. 1997. Sampling and Sample Preparation. In: Gill, R. (Ed.), *Modern Analytical Geochemistry: An Introduction to Quantitative Chemical Analysis for Earth, Environmental and Material Scientists*. Routledge: New York, USA. 342pp.
- Rao, Y.R., and D.J. Schwab. 2007. Transport and Mixing Between the Coastal and Offshore Waters in the Great Lakes: a Review. *Journal of Great Lakes Research*. Vol. 33, No. 1, pp. 202-218.
- Rasul, N., J.P. Coakley, and R. Pippert. 1997. Sedimentary Environment of Lake Erie: Geologic Setting, Sediment Distribution, and Modern Evolutionary Trends. In M. Munawar (Ed.), *State of Lake Erie: Past, Present, and Future*. Backhuys Publishers. Leiden: The Netherlands.
- Regional Planning Association (RPA), 2008. The Emerging Megaregions. America 2050. Available online: www.america2050.org/megaregions.html. [Accessed: July 16, 2017].
- Rinner, C. 2007. A Geographic Visualization Approach to multi-criteria evaluation of urban quality of life. *International Journal of Geographical Information Science*. Vol. 21, No. 8, pp. 907-919.
- Roe, A. 1953. The Making of a Scientist. Dodd, Mead & Company. New York. 244 pp. Available online: www.gwern.net/docs/iq/1953-roe-makingscientist.pdf. [Accessed: August 13, 2017].
- Ruda, A. 2016. Spatial Decision Support using Data Geovisualiation: The Example of the Conflict Between Landscape Protection and Tourism Development. *Journal of Maps*. Vol. 12, No. 5, pp. 1262-1267.
- Rukavina, N., M. Dunnett, and C. Prokopec. 2013. Great Lakes Sediment Database. Environment and Climate Change Canada. Available online: www.ec.gc.ca/inre-nwri/default.asp?lang=En&n=9890771E-1. [Accessed: July 16, 2017].
- Sabo, A., A.M. Gani, and A.Q. Ibrahim. 2013. Pollution Status of Heavy Metals in Water and Bottom Sediment of River Delimi in Jose, Nigeria. *American Journal of Environmental Protection*. Vol. 1, No. 3, pp. 47-53.

- Serafim, A., R. Company, B. Lopes, J. Rosa, A. Cavaco, G. Castela, N. Olea, and M.J. Bebianno. 2012. Assessment of Essential and Nonessential Metals and Different Metal Exposure Biomarkers in the Human Placenta in a Population from the South of Portugal. *Journal of Toxicology and Environmental Health, Part A*. Vol. 13, No. 13-15, pp. 867-877.
- Simpson, G. and Y.H. Wu. 2014. Accuracy and Effort of Interpolation and Sampling: Can GIS Help Lower Field Costs? *ISPRS International Journal of Geo-Information*. Vol. 3, No. 4, pp. 1317-1333.
- Slocum, T.A., C. Blok, B. Jiang, A. Koussoulakou, D.R. Montello, S. Fuhrmann, and N.R. Hedley. 2001. Cognitive and Usability Issues in Geovisualization. *Cartography and Geographic Information Science*. Vol. 28, No. 1, pp. 61-75.
- Soong, K.L. 1974. Versuche zur adsorptiven Bindung von Schwermetall Ionen an künstlichen Tongemischen. Unpublished PhD Thesis, University of Heidelberg, Germany. Available online: www.tandfonline.com/doi/pdf/10.1080/02626669709492068?needAccess=true. [Accessed: July 31, 2017].
- Stamper, V., C. Copeland and M. Williams. 2012. Poisoning the Great Lakes: Mercury Emissions from Coal-Fired Power Plants in the Great Lakes Region. Natural Resources Defense Council. Available online: <https://www.nrdc.org/sites/default/files/poisoning-the-great-lakes.pdf>. [Accessed: July 21, 2016].
- Stephanson, L. 2009. Environment Canada's Gasoline Regulations: Discussion Paper Proposing Indeterminate Exemption for the Use of Leaded Gasoline in Competition Vehicles. Environment and Climate Change Canada. Available online: www.ec.gc.ca/lcpe-cepa/default.asp?lang=En&n=54FE5535-1&wsdoc=8E3C2E9B-38A8-461A-8EC3-C3AA3B1FD585. [Accessed: July 16, 2017].
- Tobler, W. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*. Vol. 46, No. 2, pp. 234-240.
- Thomas, R.L., J-M. Jaquet, and A.L.W. Kemp. 1976. Surficial Sediments of Lake Erie. *Journal of the Fisheries Resource Board of Canada*. Vol. 33, No. 3, pp. 385-403.
- Tran, M. 2009. Administrative Law: The Reach of Procedural Fairness Rights [Lecture Slides]. Available online: www.yorku.ca/igreene/documents/tranAdministrativeLawChpt8.ppt. [Accessed: July 16, 2017].
- Tsuchiya, K. 1969. Causation of Ouch-Ouch Disease (Itai-Itai Byō): An Introductory Review. Part II. Epidemiology and Evaluation. *Keio Journal of Medicine*. Vol. 18, No. 4, pp. 181-194.
- U.S. Army Engineer District (USAED) Detroit, Michigan. 1974. Confined Disposal Facility at Pointe Mouillee for Detroit and Rouge Rivers: Final Environmental Statement. Available online: www.dtic.mil/docs/citations/ADA107369. [Accessed: June 4, 2017].

- U.S. Army Corps of Engineers (USACE); Great Lakes and Ohio River Division. 2012. Great Lakes System: Dredged Material Management Long Term Strategic Plan. Available online: www.lre.usace.army.mil/Portals/69/docs/Navigation/Great%20Lakes%20Dredged%20Material%20Management%20Long%20Term%20Strategic%20Plan.pdf. [Accessed: July 31, 2017].
- U.S. Department of Health and Human Services (USHHS). 2012. Toxicological Profile for Cadmium. Public Health Service: Agency for Toxic Substances and Disease Registry. Atlanta, Georgia. Available online: www.atsdr.cdc.gov/toxprofiles/tp5.pdf. [Accessed: July 16, 2017].
- U.S. Environmental Protection Agency (EPA). 1965. Report on Pollution of the Detroit River, Michigan Waters of Lake Erie, and Their Tributaries; Summary, Conclusions and Recommendations. Available online: www.books.google.ca/books/about/Report_on_Pollution_of_the_Detroit_River.html?id=DePjwEACAAJ&redir_esc=y. [Accessed: June 4, 2017].
- U.S. Environmental Protection Agency (EPA). 1996. EPA Takes Final Step in Phaseout of Leaded Gasoline. EPA Press Release. US Environmental Protection Agency. Available online: www.archive.epa.gov/epa/aboutepa/epa-takes-final-step-phaseout-leaded-gasoline.html. [Accessed: July 16, 2017].
- U.S. Environmental Protection Agency (EPA). 2015. Great Lakes Facts and Figures. US Environmental Protection Agency. Available online: www.epa.gov/greatlakes/great-lakes-facts-and-figures.html. [Accessed: June 26, 2017].
- U.S. Geological Survey (USGS). 2014. Mercury in Aquatic Ecosystems. National Water Quality Assessment (NAWQA) Project. Available online: <https://water.usgs.gov/nawqa/mercury/MercuryFAQ.html>. [Accessed: October 21, 2017].
- Verma, S., R. Bhattarai, N.S. Bosch, R.C. Cooke, P.K. Kalita, M. Markus. 2015. Climate Change Impacts on Flow, Sediment and Nutrient Export in a Great Lakes Watershed Using SWAT. *Clean – Soil, Air, Water (CSAWAC)*. Vol. 43, No. 11, pp. 1445-1558.
- Verly, G, M. David, A.G. Journel, and A. Marechal. 1983. Geostatistics for Natural Resource Characterization, Part 1. In *Proceedings of the NATO Advanced Study Institute on Geostatistics for Natural Resource Characterization*. South Lake Tahoe, California, USA, September 6-17, 1983.
- Villard, M-A. and B.A. Maurer. 1996. Geostatistics as a Tool for Examining Hypothesized Declines in Migratory Songbirds. *Ecology*. Vol. 77, No. 1, pp. 59-68.
- Watanabe, K., E. Kobayashi, Y. Suwazono, Y. Okubo, T. Kido, K. Nogawa. 2004. Tolerable Lifetime Cadmium Intake Calculated from the Inhabitants Living in the Jinzu River Basin, Japan. *Bulletin of Environmental Contamination and Toxicology*. Vol. 72, No. 6, pp. 1091-1097.
- Webster, R. and M.A. Oliver. 1993. How Large a Sample is needed to Estimate the Regional Variogram Adequately? *Geostatistics Troia '92*. Vol. 1, pp. 155-166.

- Webster, R. and M.A. Oliver. 2007. Geostatistics for Environmental Scientists (Second Edition). John Wiley & Sons, Ltd.: West Sussex, England. 330pp.
- Wester, R.C., H.I. Maibach, L. Sedik, J. Melendres, S. DiZio, and M. Wade. 1992. In Vitro Percutaneous Absorption of Cadmium from Water and Soil into Human Skin. *Fundamental and Applied Toxicology*. Vol. 19, No. 1, pp. 1-5.
- Williams, J., M.A. Arsenault, B.J. Buczkowski, J.A. Reid, J.G. Flocks, M.A. Kulp, S. Penland, and C.J. Jenkins. 2011. Surficial Sediment Character of the Louisiana Offshore Continental Shelf Region: A GIS Compilation. U.S. Geological Survey Open-file Report 2006-1195. Available online: www.pubs.er.usgs.gov/publication/ofr20061195. [Accessed: July 16, 2017].
- Willmott, C.J. 1984. On the Evaluation of Model Performance in Physical Geography. In G.L. Gaile and C.J. Willmott (Ed.), *Spatial Statistics and Models*. D. Reidel Publishing Company. Dordrecht: Holland.
- Wu, C., J. Wu, Y. Luo, H. Zhang, Y. Teng, & S. D. DeGloria. 2011. Spatial interpolation of severely skewed data with several peak values by the approach integrating kriging and triangular irregular network interpolation. *Environmental Earth Sciences*. Vol. 63, No. 5, pp. 1093-1103.
- Wuana, R.A. and F.E. Okieimen. 2011. Heavy Metals in Contaminated Soils: A Review of Sources, Chemistry, Risks and Best Available Strategies for Remediation. *International Scholarly Research Network: Ecology*. Vol. 2011, pp. 1-20.
- Yan, L., S. Zhou, W. Ci-fang, L. Hong-yi, and L. Feng. 2007. Improved Prediction and Reduction of Sampling Density for Soil Salinity by Different Geostatistical Methods. *Agricultural Sciences in China*. Vol. 6, No. 7, pp. 832-841.
- Zaidi, A., P.A. Wani, and M.S. Khan. 2012. Toxicity of Heavy Metals to Legumes and Bioremediation. Springer-Verlag Wien. Pondicherry, India. 248 pp.
- Ziary, Y., and H. Safari. 2007. To Compare Two Interpolation Methods: IDW, Kriging for Providing Properties (Area) Surface Interpolation Map Land Price. District 5, Municipality of Tehran Area 1. In *Proceedings of the Strategic Integration of Surveying Services, FIG Working Week*. Hong Kong SAR, Chin, May 13-17, 2007.