# Non-Halogenated and Halogenated Heterocyclic Aromatic Compounds in the Canadian Environment

by

Nipuni N. Vitharana

A thesis submitted to the Faculty of Graduate and Postdoctoral Studies of the
University of Manitoba
in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry
University of Manitoba
Winnipeg, Manitoba, Canada

## ABSTRACT

Polycyclic aromatic compounds (PACs) are a broad class of pollutants with two or more fused aromatic rings, originating from petrogenic, pyrogenic, and biogenic sources. Hetero-polycyclic aromatic compounds (HPACs), a subclass of PACs that contains nitrogen, sulfur, and/or oxygen in their aromatic structure, are understudied despite growing evidence of their persistence, bioaccumulation, and toxicity. This is largely because the United States Environmental Protection Agency and Canada's Environmental Protection Act, 1999, recognize only 16 polycyclic aromatic hydrocarbons (PAHs) as priority pollutants, leaving other subclasses of PACs underrepresented in environmental monitoring. Consequently, my research was geared toward addressing the paucity of scientific data on HPACs. In doing so, I generated new knowledge on HPACs across several fronts. First, I developed and validated both targeted and non-targeted analytical methods for accurately measuring HPACs in environmental matrices. These methods were then applied to generate baseline concentration data for a broad suite of PACs, including 50 HPACs in Lake Erie biota, and Lake Ontario sediments, many of which had not previously been reported in these systems. Second, using advanced chemometric techniques, such as multivariate and cluster analysis, I demonstrated that HPACs have greater diagnostic power in deriving the sources, fate, and behaviour of PACs relative to their homocyclic counterparts. Notably, comparison of HPAC profiles in Lake Ontario sediments against coal-tar- and non-coal-tar-based reference materials revealed that Hamilton Harbour sediments are primarily contaminated by coal-tar-derived PAC sources, illustrating the real-world diagnostic value of HPACs. Third, I showed that HPAC distribution patterns clearly differentiated fish communities by habitat and physiological traits, such as biotransformation capacity and body size, whereas homocyclic PAC profiles were comparatively uniform, indicating weaker habitat-specific influences. Fourth, I developed a multi-

tiered data-filtering framework to help the scientific community streamline the interpretation of high-resolution mass spectra to elucidate PACs in complex environmental matrices. Taken together, my research demonstrates the importance of this overlooked class of contaminants and highlights the need for the inclusion of HPACs into long-term monitoring and risk assessment frameworks, which will ultimately improve our understanding of the environmental behaviour and impacts of PACs across Canadian environmental systems.

## PREFACE

My thesis is an original work entirely written by Nipuni Vitharana. Due to the style of the "sandwich method", there might be some redundancies in the introductory paragraphs of the six manuscripts as well as the introductory chapter.

Chapter 3 of my thesis is a manuscript published in the *Journal of Chromatography A*. I was responsible for performing the analyses, processing the data, and writing this manuscript. Thor Halldorson helped with method development and conceptualization. Dr. Zhe Xia, Olga Francisco, Erin Liebzeit, Amica-Mariae Lucas, and Vida Moradi helped with the methodology. Drs. Chris Marvin and Philippe J. Thomas reviewed the document and helped draft the manuscript. Dr. Gregg Tomy, as the primary investigator of this study, was responsible for the acquisition of the funds, project administration, supervision, experiment conceptualization, and manuscript review.

Chapter 4 of my thesis is a manuscript published in *Science of the Total Environment*. I was responsible for performing the analyses, processing the data, conducting statistical analysis, and writing this manuscript. Thor Halldorson helped with method development and conceptualization. Mike Dereviankin helped with the statistical analysis. Brett Paddock and Aviva Gerring helped with methodology. Drs. Chris Marvin and Paul Helm provided environmental samples, reviewed the document, and helped draft the document. Drs. Philippe J. Thomas and Ifeoluwa Grace Idowu reviewed the document. Dr. Gregg Tomy, as the primary investigator of this study, was responsible for the acquisition of the funds, project administration, supervision, experiment conceptualization, and manuscript review.

Chapter 5 of my thesis is a manuscript published in *Environmental Pollution*. I was responsible for performing the analyses, processing the data, conducting statistical analysis, and writing this

manuscript. Thor Halldorson helped with method development and conceptualization. Brett Paddock, Thane Tomy, and Aviva Gerring helped with methodology. Drs. C.E. Heuvel and A.T. Fisk provided environmental samples, reviewed the document, and helped draft the document. Dr. Chris Marvin reviewed the document. Dr. Gregg Tomy, as the primary investigator of this study, was responsible for the acquisition of the funds, project administration, supervision, experiment conceptualization, and manuscript review.

Chapter 6 of my thesis is a manuscript published in the *Journal of Hazardous Materials*. I was responsible for performing the analyses, processing the data, conducting statistical analysis, and writing this manuscript. Thor Halldorson helped with method development and conceptualization. Brett Paddock and Thane Tomy helped with methodology. Drs. C.E. Heuvel and A.T. Fisk provided environmental samples, reviewed the document, and helped draft the document. Dr. Chris Marvin reviewed the document. Dr. Gregg Tomy, as the primary investigator of this study, was responsible for the acquisition of the funds, project administration, supervision, experiment conceptualization, and manuscript review.

Chapter 7 of my thesis is a viewpoint published in *Environmental Science and Technology*. I was responsible for writing the document. Drs. Christopher Marvin, Philippe J. Thomas, Ifeoluwa Grace Idowu, Courtney D. Sandau, Reyd Dupuis-Smith, Jennifer F. Provencher, Aaron T. Fisk, and Gregg T. Tomy reviewed the document.

Chapter 8 of my thesis is a manuscript that is to be submitted to *Environmental Science and Technology*. I was responsible for performing the analyses, processing the data, and writing this manuscript. Thor Halldorson helped with method development and conceptualization. Brett Paddock and Thane Tomy helped with methodology. Drs. Chris Marvin, C.E. Heuvel, and A.T.

Fisk provided environmental samples, reviewed the document, and helped draft the document. Dr. Bernard Crimmins helped draft and review the document. Dr. Gregg Tomy, as the primary investigator of this study, was responsible for the acquisition of the funds, project administration, supervision, experiment conceptualization, and manuscript

## LIST OF PUBLICATIONS

1. **Vitharana, Nipuni N**; Halldorson, Thor; Paddock, Brett; Tomy, Thane; Gerring, Aviva; Marvin, Chris; Heuvel, CE; Fisk, AT; Crimmins, Bernard.; Tomy, Gregg T.; Comprehensive Data Filtering Approaches for the Non-Targeted Analysis of Homocyclic, Heterocyclic and Halogenated Polycyclic Aromatic Compounds (To be submitted to *Environmental Science & Technology*, **2026**)

2. **Vitharana, Nipuni N**; Marvin, Christopher; Thomas, Philippe J; Idowu, Ifeoluwa Grace; Sandau, Courtney D; Dupuis-Smith, Reyd; Provencher, Jennifer F; Fisk, Aaron T; Tomy, Gregg T; Heterocyclic Aromatic Compounds: The Next Frontier in Environmental Forensic Science, *Environmental Science & Technology*, 2026.

3. **Vitharana, Nipuni N**; Halldorson, Thor; Paddock, Brett; Tomy, Thane; Gerring, Aviva; Marvin, Chris; Heuvel, CE; Fisk, AT; Tomy, Gregg T; Using Chemometrics to Discern Patterns and Spatial Distribution of Heterocyclic Aromatic Compounds in Lake Erie Fish, *Environmental Pollution*, 2026, 128121

4. **Vitharana, Nipuni N**; Halldorson, Thor; Paddock, Brett; Tomy, Thane; Marvin, Chris; Fisk, AT; Heuvel, CE; Tomy, Gregg T; Habitat and Diet Affect Bioaccumulation of Homocyclic and Heterocyclic Aromatic Compounds in Lake Erie Fish, *Journal of Hazardous Materials*, **2026**, 142044

5. Liebzeit, Erin C; **Vitharana, Nipuni N**; Halldorson, Thor; Provencher, Jennifer F; Mallory, Mark L; Tomy, Gregg T; Dupuis-Smith, Reyd; Species-and Tissue-Specific Accumulation of Polycyclic Aromatic Compounds in Three Arctic Seabirds, *Environmental Science & Technology*, **2026**, volume 60-10, 8062-8073.

6. Xia, Zhe; Halldorson, Thor; **Vitharana, Nipuni**; Kim, Marcus; Daramola, Dami; Marvin, Chris; Thomas, Philippe J; Dupuis-Smith, Reyd A; Provencher, Jennifer F; Tomy, Gregg T; Dynamic multiple reaction monitoring for high throughput detection and quantitation of polycyclic aromatic compounds*, Analytical Methods*, **2026**, volume 18-7, 1503-1512.

7. Xia, Zhe; Gao, Xinyu; Halldorson, Thor; **Vitharana, Nipuni**; Marvin, Chris; Thomas, Philippe J; Tomy, Gregg T; An Efficient and Greener Alternative for the Extraction of Polycyclic Aromatic Compounds from Sediments, *Separations*, **2026**, volume 13-2, 68.

8. Idowu, Ifeoluwa Grace; Dereviankin, Mike; Halldorson, Thor; **Vitharana, Nipuni**; Harbicht, Lacey; Helm, Paul; Kleywegt, Sonya; Marvin, Chris; Sandau, Courtney D; Tomy, Gregg T; Source characterization of polycyclic aromatic hydrocarbons in urban sediments: Insights from the contribution of coal tar, sealants, and other related industrial materials in Lake Ontario, *Science of The Total Environment*, **2025**, volume 1009,18096.

9. Dubiel, Justin; Mertens, Emily; Boyte, Vann; Eagon, Lauren; **Vitharana, Nipuni**; Gao, Xinyu; Hartmann, Hannah; Terrazas-Salgado, Luis; Xia, Zhe; Tomy, Gregg; Chlorination and Bromination of Anthracene Affects Aryl Hydrocarbon Receptor Activation and Early Life Stage Mortality in Zebrafish (Danio rerio, *Environmental Science & Technology*, **2025**, volume 59-28, 14347-14358.

10. **Vitharana, Nipuni N**; Halldorson, Thor; Dereviankin, Mike; Paddock, Brett; Gerring, Aviva; Marvin, Chris; Helm, Paul; Thomas, Philippe J; Idowu, Ifeoluwa Grace; Tomy, Gregg T; Heterocyclic aromatic compounds in Lake Ontario sediments: Spatiotemporal distribution and new approaches to fingerprinting, *Science of The Total Environment*, **2025**, 986, 179802.

11. Zahaby, Yasmeen; Crump, Doug; O'Brien, Jason; Dupuis-Smith, Reyd; Dwyer-Samuel, Frederic; Laing, Rodd; Pilgrim, Samantha; Gear, George; Pamak, Carla; Saunders, Michelle; Comparison of gene expression and polycyclic aromatic compound profiles in hepatic tissue of black guillemot (Cepphus grylle) collected from an oil spill site and a non-spill site in the Arctic, *Marine Pollution Bulletin*, **2025**, 212, 117504.

12. Loudon, Kara B; Tomy, Thane MZ; Liebzeit, Erin C; Halldorson, Thor; Xia, Zhe; Sambanthan, Sara; Hoang, Duc Luong; **Vitharana, Nipuni**; Tomy, Gregg T; Estimating dietary exposure to polycyclic aromatic compounds from food grade plastics, *Environmental Science: Advances*, 2025, volume 4-2, 292-305.

13. **Vitharana, Nipuni N**; Halldorson, Thor; Xia, Zhe; Francisco, Olga; Marvin, Chris; Thomas, Philippe J; Liebzeit, Erin; Lucas, Amica-Mariae; Moradi, Vida; Tomy, Gregg T; A validated approach for analysis of heterocyclic aromatic compounds in sediment samples, *Journal of Chromatography A*, **2024**, volume 1718, 464723.

14. Moradi, Vida; Halldorson, Thor; Xia, Zhe; **Vitharana, Nipuni**; Marvin, Chris; Thomas, Philippe J; Sorais, Manon; Crossin, Glenn T; Tomy, Gregg; Determining Polycyclic Aromatic Compounds in Bird Feathers Using Pressurized Fluid Extraction, *Separations*, **2023,** volume 10-9, 503.

15. Moradi, Vida; Halldorson, Thor; Idowu, Ifeoluwa; Xia, Zhe; **Vitharana, Nipuni**; Marvin, Chris; Thomas, Philippe J; Tomy, Gregg T; Microbead-Beating Extraction of Polycyclic Aromatic Compounds from Seabird Plasma and Whole Blood, *Separations*, **2023,** volume 10-1, 48.

## ACKNOWLEDGEMENTS

This thesis would not have been possible without the support and guidance of many people.

I want to thank my advisor, Gregg Tomy. His continuous guidance and support over the past four years have shaped me into a better scientist and a better person. I truly appreciate the support he provided my husband and me to survive in a country thousands of miles from our families. Beyond the science, the most important life lesson I learned from Gregg is humility, a lesson I will carry with me throughout my professional life.

I am very appreciative of the suggestions, contributions, constructive criticism, and advice from my Ph.D. committee members, Drs. Sheryl Tittlemier, Mark Fry, and Jorg Stetefeld. I also want to thank the external examiner of this thesis, Dr. Hui Peng for his time and effort.

I want to thank my 'go-to' person whenever I was in trouble in the lab, Thor Halldorson. He ensured I had the scientific knowledge and experience needed to complete this thesis, and I am grateful for all his guidance from the start of my program. I want to thank both Gregg and Thor for being the best two teachers I have met in my 20 years long student life. Furthermore, I want to thank all other current and former members and graduate students of the Center of Oil and Gas Research and Development (COGRAD): Dr. Ifeoluwa Grace Idowu, Olga Francisco, Dr. Zhe Xia, Sara Sambanthan, Vida Moradi, Erin Liebzeit, Sithumi Yaddehige, Jordon Wong, and Ayesha Evelyn for their support in my research. Special thanks to Dr. Grace Idowu, who taught me how to use the high-resolution mass spectrometer system and continues to do so even today. Sara and Vida, thank you for making my first year here so easy.

I would also like to thank all the summer students, especially Brett Paddock, Thane Tomy, Aviva Gerring, Kara Laudon, and Amica Mariae Lucus, who helped me process the samples used in this

study. I am especially appreciative of the Department of Chemistry at the University of Manitoba for its assistance throughout my program. I also thank the University of Manitoba Graduate Fellowship (UMGF) for the financial support.

I am particularly grateful to Dr. Chris Marvin for the many fruitful conversations we have had together throughout my journey as a PhD student. I also want to thank Drs. Aaron Fisk, Chris Marvin, Paul Helm, Philippe Thomas, and Cecilia Heuvel for providing the environmental samples and reference materials used in this study. I also thank Mike Dereviankin for providing me with initial guidance on chemometric analysis.

Next, I want to thank my favorite people in the world, my family. First, I really appreciate my husband, Dilan Witharanage, who made sure I was loved, fed, driven to the lab, and that my tuition was paid on time so I could focus on this program without any trouble. Without you by my side, any of these would not have been possible. Second, I want to thank my parents, Pamoda Basnayaka and Sarath Vitharana, who always encouraged me to pursue my dreams and become the best version of myself. I also want to thank my younger sisters, Kaushalya and Kavya, for always being there for me and being my best friends. Special thanks go to Amma (mom), Kaushi, and Kavya for keeping my anxiety in check through their long video calls.

Also, I want to thank my in-laws, Dilan's family. I would like to thank his parents, Chandani Amarakoon and Sumanasiri Witharanage, and his sister, Gayashani Witharanage. From the first day I knew them, they were nothing but supportive of me. I always appreciate Dilan's mom's encouraging and loving words. I also want to thank all the relatives and friends who helped Dilan's mom and dad during my father-in-law's health issues. Special thanks go to Amma, Kaushi, Thaththa, Punchi (Kamali Basnayaka), Chuti akka (Achala Gamaethige), Ranga ayya, and Isuru

ayya, and my friends, Udari and Navoda (and Gayan, of course), for their support during the hardest time of these four years.

I also want to thank all my friends who helped me over the last four years. I want to thank my best friends, Sachini and Jayamal, for coming to Manitoba and becoming our family away from family. In 2018, I never imagined our double dates would still be happening in Canada years later. I would like to thank my school best friends, Dileesha, Navoda, and Sandarashmi, for making me a part of something special like “Sanidina Rasakatha”. I also want to thank my friends from the University of Sri Jayewardenepura: Nishadi, Lashan, and Imi, who supported me greatly throughout my PhD application process. I also thank my darling Yasandi for making sure I was doing well, even being thousands of miles away. Also, thank you for all the group chats that made me feel closer to home.

I thank everyone who helped us settle in Canada. Especially, I want to thank Geethani akka and Sumudu ayya for all the help you gave us.

Finally, I want to thank all my teachers, from grade one through my undergraduate studies, for uncovering this path for me. I appreciate all the teachers I had from grades one to five at Kalawana National School, Sri Lanka, and all my teachers from Sumana Balika Vidyalaya, Ratnapura, Sri Lanka. I want to thank my advanced-level Chemistry teachers, Ms. Chamila Dharmadasa, Mr. Laal Diyamullage (Laal loku thaththa), and Mr. Harshana Chamal. I am grateful to all the lecturers in the Department of Chemistry at the University of Sri Jayewardenepura, Sri Lanka. I want to thank my undergraduate supervisor, Prof. S. P. Deraniyagala, and my two co-supervisors, Prof. Theshini Perera and Dr. Asitha Cooray. I am always grateful for the guidance and support you three have provided me during my undergraduate studies.

## DEDICATION

To my husband, Sanhka: my soulmate, my best friend, and my greatest source of strength and happiness.

# TABLE OF CONTENTS

**ABSTRACT** ..... **II**

**PREFACE** ..... **IV**

**LIST OF PUBLICATIONS** ..... **VII**

**ACKNOWLEDGEMENTS** ..... **IX**

**DEDICATION** ..... **XII**

**TABLE OF CONTENTS** ..... **XIII**

**LIST OF FIGURES** ..... **XIX**

**LIST OF TABLES** ..... **XXII**

**APPENDIX** ..... **XXIII**

**LIST OF ABBREVIATIONS** ..... **XXVIII**

**CHAPTER 1 INTRODUCTION** ..... **1**

**1.1 POLYCYCLIC AROMATIC COMPOUNDS** ..... **1**

**1.2 HETERO-POLYCYCLIC AROMATIC COMPOUNDS** ..... **5**

**1.3 ELECTRONIC AND PHYSICOCHEMICAL PROPERTIES OF HETERO-POLYCYCLIC AROMATIC COMPOUNDS** ..... **8**

**1.4. ENVIRONMENTAL SOURCES OF HETEROCYCLIC AROMATIC COMPOUNDS** ..... **16**

**1.5 ENVIRONMENTAL DISTRIBUTION OF HETERO-POLYCYCLIC AROMATIC COMPOUNDS** ..... **18**

**1.6 NEGATIVE IMPACTS OF HETERO-POLYCYCLIC AROMATIC COMPOUNDS ON HUMANS AND OTHER BIOGENIC ENVIRONMENT** ..... **26**

1.6.1 Toxic properties of hetero-polycyclic aromatic compounds ..... 26

**1.7 SCOPE OF MY RESEARCH ..... 29**

**1.8 RESEARCH HYPOTHESES ..... 29**

**1.9 THESIS STRUCTURE ..... 30**

**1.10 REFERENCES ..... 31**

**CHAPTER 2 METHODOLOGY ..... 45**

**2.1 SAMPLE PREPARATION ..... 45**
2.1.1 Sample collection and pre-preparation ..... 45
2.1.2 Extraction and purification ..... 48

**2.2 DATA ACQUISITION AND QUANTIFICATION ..... 53**
2.2.1 Gas chromatography ..... 54
2.2.2 Mass spectrometry ..... 55
2.2.3 Gas chromatography coupled with tandem mass spectrometry ..... 58
2.2.4 Two-dimensional gas chromatography coupled with high resolution time of flight mass spectrometry ..... 66

**2.3 DATA PROCESSING AND INTERPRETATION ..... 70**
2.3.1 Statistical analysis ..... 71
2.3.2 Non-targeted data analysis for the identification of polycyclic aromatic compounds ..... 81

**2.4: REFERENCES ..... 82**

**CHAPTER 3 A VALIDATED APPROACH FOR ANALYSIS OF HETEROCYCLIC AROMATIC COMPOUNDS IN SEDIMENT SAMPLES ..... 92**

**3.1 ABSTRACT ..... 93**

**3.2 INTRODUCTION ..... 94**

**3.3 MATERIALS AND METHODS ..... 96**
3.3.1 Chemicals ..... 96
3.2.2 Environment sample ..... 96
3.2.2 Sample extraction and cleanup ..... 97
3.2.3 Gas chromatography tandem mass spectrometry ..... 98
3.2.4 Method validation ..... 104

**3.3 RESULTS AND DISCUSSION ..... 109**
3.3.1 Mass spectral characterization ..... 109

3.3.2 Extraction optimization ..... 110
3.3.3 Performance characteristics ..... 113

**3.4 CONCLUSIONS ..... 116**

**3.5 ACKNOWLEDGEMENTS ..... 116**

**3.6 REFERENCES ..... 117**

**CHAPTER 4 HETEROCYCLIC AROMATIC COMPOUNDS IN LAKE ONTARIO SEDIMENTS: SPATIOTEMPORAL DISTRIBUTION AND NEW APPROACHES TO FINGERPRINTING ..... 121**

**4.1 ABSTRACT ..... 122**

**4.2 INTRODUCTION ..... 123**

**4.3 MATERIALS AND METHOD ..... 125**
4.3.1 Chemicals ..... 125
4.3.2 Reference materials ..... 125
4.3.3 Sediment samples ..... 126
4.3.4 Extraction and cleanup procedures ..... 126
4.3.5 Gas chromatography mass spectrometry ..... 127
4.3.6 Quality control ..... 128
4.3.7 Fingerprinting analysis using principal component analysis and uniform manifold approximation and projection ..... 129

**4.4 RESULTS AND DISCUSSION ..... 130**
4.4.1 Spatial distribution of HPACs ..... 130
4.4.2 Fingerprinting analysis of HPACs ..... 136

**4.5 CONCLUSIONS ..... 146**

**4.6 ACKNOWLEDGEMENTS ..... 147**

**4.7 REFERENCES ..... 148**

**CHAPTER 5 USING CHEMOMETRICS TO DISCERN PATTERNS AND SPATIAL DISTRIBUTION OF HETEROCYCLIC AROMATIC COMPOUNDS IN LAKE ERIE FISH ..... 152**

**5.1 ABSTRACT ..... 153**

**5.2. INTRODUCTION ..... 154**

**5.3. MATERIALS AND METHODS ..... 156**
5.3.1 Chemicals ..... 156
5.3.2 Biota samples ..... 157
5.3.3 Extraction and cleanup procedures ..... 157
5.3.4 Gas chromatography tandem mass spectrometry ..... 158
5.3.5 Method validation ..... 159
5.3.6 Statistical analysis ..... 159

**5.4. RESULTS AND DISCUSSION ..... 161**

**5.4.1. Profiles and concentrations of polycyclic aromatic compounds in biota ..... 161**
5.4.2 Spatial distribution of PACs in Lake Erie biota ..... 170

**5.5 CONCLUSIONS ..... 181**

**5.6 ACKNOWLEDGEMENTS ..... 182**

**5.7 DECLARATION OF GENERATIVE AI AND AI ASSISTED TECHNOLOGIES IN THE MANUSCRIPT PREPARATION PROCESS ..... 182**

**5.8 REFERENCES ..... 183**

**CHAPTER 6 : HABITAT AND DIET AFFECT BIOACCUMULATION OF HOMOCYCLIC AND HETEROCYCLIC AROMATIC COMPOUNDS IN LAKE ERIE FISH ..... 188**

**6.1 ABSTRACT ..... 189**

**6.2 ENVIRONMENTAL IMPLICATIONS ..... 190**

**6.3 INTRODUCTION ..... 191**

**6.4 MATERIALS AND METHODS ..... 193**
6.4.1 Chemicals ..... 193
6.4.2 Biota samples collection and stable isotope analysis ..... 193
6.4.3 Extraction, cleanup and gas chromatography tandem mass spectrometry analysis ..... 194
6.4.4 Multivariate analysis ..... 195
6.4.5 Quantification of biomagnification ..... 197

**6.5. RESULTS AND DISCUSSION ..... 198**
6.5.1 Profiles and concentrations of polycyclic aromatic compounds in biota ..... 198
6.5.2 Species and habitat driven clustering based on multivariate analysis ..... 204

6.5.3 Bioaccumulation of polycyclic aromatic compounds in Lake Erie food web ..... 212
6.5.4 Future management of PACs in the Lake Erie ecosystem ..... 215

**6.6 CONCLUSIONS ..... 217**

**6.7 ACKNOWLEDGEMENTS ..... 218**

**6.8 DECLARATION OF GENERATIVE AI AND AI ASSISTED TECHNOLOGIES IN THE MANUSCRIPT PREPARATION PROCESS ..... 218**

**6.9 REFERENCES ..... 219**

**CHAPTER 7 HETEROCYCLIC AROMATIC COMPOUNDS: THE NEXT FRONTIER IN ENVIRONMENTAL FORENSIC SCIENCE ..... 224**

**7.1 A VIEWPOINT ON HPACs ..... 225**

**7.2 ACKNOWLEDGEMENTS ..... 229**

**7.3 REFERENCES ..... 230**

**CHAPTER 8 COMPREHENSIVE DATA FILTERING APPROACHES FOR NON-TARGETED ANALYSIS OF POLYCYCLIC AROMATIC COMPOUNDS ..... 232**

**8.1 ABSTRACT ..... 233**

**8.2. INTRODUCTION ..... 234**

**8.3. MATERIALS AND METHODS ..... 234**
8.3.1 Chemicals ..... 236
8.3.2 Environmental samples ..... 237
8.3.3 Sample purification and extraction ..... 237
8.3.4 Gas chromatography mass tandem spectrometry analysis ..... 237
8.3.5 GC x GC high resolution time of flight mass spectrometry analysis ..... 238
8.3.6 Data analysis ..... 239

**8.4. DISCUSSION ..... 242**
8.4.1. Preliminary data filtering through ChromaTOF ..... 242
8.4.3. Workflow 1: assignment of confidence levels based on a custom-created library data of PACs ..... 243
8.4.4. Workflow 2: identification of potential halogenated PACs ..... 243
8.4.5. Workflow 3: identification of potential heterocyclic aromatic compounds ..... 258
8.4.6. Workflow 4: results integration and PAC identification ..... 258

8.4.7. Targeted and non-targeted PACs were detected in the environmental samples ..... 262
8.4.8. Rationale for Developing a Novel Workflow..... 262

**8.5 CHALLENGES AND LIMITATIONS OF THE DATA ANALYSIS PROCEDURE ... 262**

**8.6 FUTURE DIRECTIONS..... 267**

**8.7 ACKNOWLEDGMENTS ..... 269**

**8.8 REFERENCES..... 270**

**CHAPTER 9: CONCLUSIONS AND FUTURE DIRECTIONS ..... 274**

**9.1 CONCLUSIONS ..... 274**

**9.2 FUTURE DIRECTIONS..... 277**

**APPENDIX I. ADDITIONAL METRICS FROM CHAPTER 3: A VALIDATED APPROACH FOR ANALYSIS OF HETEROCYCLIC AROMATIC COMPOUNDS IN SEDIMENT SAMPLES .....I**

**APPENDIX II. ADDITIONAL METRICS FROM CHAPTER 4: HETEROCYCLIC AROMATIC COMPOUNDS IN LAKE ONTARIO SEDIMENTS: SPATIOTEMPORAL DISTRIBUTION AND NEW APPROACHES TO FINGERPRINTING ..... XVII**

**APPENDIX III. ADDITIONAL METRICS FROM CHAPTER 5: USING CHEMOMETRICS TO DISCERN PATTERNS AND SPATIAL DISTRIBUTION OF HETEROCYCLIC AROMATIC COMPOUNDS IN LAKE ERIE FISH ..... XXXIX**

**APPENDIX IV. ADDITIONAL METRICS FROM CHAPTER 6: HABITAT AND DIET AFFECT BIOACCUMULATION OF HOMOCYCLIC AND HETEROCYCLIC AROMATIC COMPOUNDS IN LAKE ERIE FISH ..... LVIII**

**APPENDIX V: ADDITIONAL METRICS FROM CHAPTER 8: COMPREHENSIVE DATA FILTERING APPROACHES FOR NON-TARGETED ANALYSIS OF POLYCYCLIC AROMATIC COMPOUNDS..... LXXXIII**

# LIST OF FIGURES

**Figure 1.1:** The chemical structures of the United States Environmental Protection Agency's 16 priority polycyclic aromatic hydrocarbons. ..... 3

**Figure 1.2:** The chemical structures of some common unsubstituted hetero-polycyclic aromatic compounds ..... 7

**Figure 2.1:** Schematic diagram of an electron ionization source used in mass spectrometry ..... 56

**Figure 2.2:** The three operational modes: full scan, product ion monitoring and multiple reaction monitoring used in GC-MS/MS in this study ..... 60

**Figure 2.3:** Schematic diagram of a two-dimensional gas chromatography setup ..... 67

**Figure 2.4:** Schematic diagram of a high-resolution time-of-flight mass spectrometer ..... 69

**Figure 2.5:** A diagram that summarizes how chemometrics are interlinked with different subject areas ..... 72

**Figure 3.1:** The MRM ion chromatogram of our HPACs analytical standard (200 pg/µL). Top panel shows analytes that elute from $r_t$: 0-14mins and bottom panel shows analytes that elute from $r_t$: 14-33 min. ISTD4: anthracene-$d_{10}$, $r_t$: 9.76 min, coelutes with N9 and N10 and is not visible in top panel. (Peaks labeled $a^*$, $b^*$ and $c^*$ in bottom panel are mass labeled compounds present in our commercial mixture that were not used in our analytical method.)..... 102

**Figure 3.2:** The MRM ion chromatogram of our halo-HPACs analytical standard (200 pg/µL). Top panel shows analytes that elute from $r_t$: 0-14mins and bottom panel shows analytes that elute from $r_t$: 14-33 min. ISTD1: naphthalene-d8, $r_t$: 4.92 min, coelutes with HH1 and is not visible in the top panel. (Peaks labeled $a^*$ and $b^*$ in bottom panel are mass labeled compounds present in our commercial mixture that were not used in our analytical method.)..... 103

**Figure 3.3:** Residual plots of quinoline using (i) the linear regression model and (ii) the average $r_f$ model.....113

**Figure 3.4:** Blank corrected concentrations (µg/g, dry weight, *n=3*) of PACs and their relative distributions in sediments from Randal Reef (Hamilton Harbour, ON, Canada). .....115

**Figure 4.1:** Concentrations of individual PANHs (ng/g, dry weight) in (a) the HH suspended sediments from 2014 with indole (top left panel), (b) without indole (top right panel), (c) the HH

and TH bottom sediments from 2014 (bottom left panel) and (d) the non-harbour bottom sediments from 2014 and 2018 (bottom right panel). ..... 134
**Figure 4.2:** Concentrations of individual PASH and PAOH (ng/g, dry weight) in (a) the HH suspended sediments from 2014 (top panel), (b) the HH and TH bottom sediments from 2014 (middle panel) and (c) the non-harbour bottom sediments from 2014 and 2018 (bottom panel). ..... 135
**Figure 4.3:** Evaluation of chemical persistence determined from the slopes (x-axis) of the log transformed concentration of each analyte vs sampling year for suspended sediments collected from Hamilton Harbour. The dotted red lines represent the confidence interval (at 95% confidence level) derived from the average slope of benz[c]acridine, 11H-benzo[a]carbazole, dibenzothiophene, dibenzofuran and benzo[b]naphtho[2,3-d]furan (shown in yellow)..... 138
**Figure 4.4:** PAC and UMAP results between (a) PC1 and PC2 (top left panel), (b) PC1 and PC3 (top right panel) (c) variable 1 and variable 2 (bottom left panel), (d) variable 1 and variable 3 (bottom right panel) for HH harbour suspended sediments (green dots), HH harbour bottom sediments (gray dots), TH harbour Bottom sediments (red dots), non-harbour bottom sediments (purple dots), the CT reference materials (yellow dots), the NCT reference materials (orange dots) and the Mixed RM (dark blue dots). ..... 145
**Figure 5.1:** Log scaled boxplots of concentrations of major PAC classes (ng/ g, lipid-normalized) in four fish species from Lake Erie: (a) Freshwater Drum (top left), (b) Gizzard Shad (top right), (c) Yellow Perch (bottom left), and (d) Walleye (bottom right). Data are grouped by location (Eastern, East Central, West Central, and Western). ..... 170
**Figure 5.2:** Multivariate analysis of PAC profiles using principal component analysis (PCA) and uniform manifold approximation and projection (UMAP). PCA score plots are shown for (a) PC1 vs. PC2 (top left) and (b) PC1 vs. PC3 (top right), while UMAP projections are shown for (c) variable 1 vs. variable 2 (bottom left) and (d) variable 1 vs. variable 3 (bottom right). Symbols represent fish species, and colours indicate sampling regions (Eastern, East Central, West Central, and Western Lake Erie). (panh1: quinoline, panh11: acridine, panh12: 9-methylcarbazole, panh13: carbazole, panh15: 2-methylacridine, panh16: 3,6-dimethylcarbazole, panh2: isoquinoline, panh20: 11H-benzo[a]carbazole, panh21: dibenz[c,h]acridine, panh25: 5-chloroindole, panh26: 3-bromoisoquinoline, panh5: 3-methylisoquinoline, pash18:2,3-dibromobenzothiophene, pash2:

2,3-dimethylbenzothiophene, pash3:2,3,4-trimethylbenzothiophene, pash4: dibenzothiophene, pash5: 4-methyldibenzothiophene, pash6: 2,4-dimethyldibenzothiophene). ..... 176
**Figure 5.3:** Heatmaps generated from hierarchical cluster analysis (HCA) results of (a) Freshwater Drum (top left), (b) Gizzard Shad (top right), (c) Yellow Perch (bottom left), (d) Walleye (bottom right)..... 178
**Figure 6.1:** Box and whisker plots of log-transformed concentrations (lipid-corrected) of different PAC types in all the species of interest (pelagic species: purple labels and benthic species: brown labels) in order of increasing trophic positions from the western basin of Lake Erie. Species are labeled as dreissenid mussels (DM), gizzard shad (GS), spottail shiner (SS), emerald shiner (ES), yellow perch (YP), white bass (WB), walleye (WE), round goby (RG), whitesucker (WS) and freshwater drum (FD). ..... 198
**Figure 6.2:** PCA biplots between (a) PC1 and PC2 (top panel), UMAP plots between (b) variable 1 and variable 2 (middle panel-horizontal), and RDA plot between (c) RDA1 and RDA2 (bottom panel). The left panels are for the HPACs-based analyses, the middle panels (vertical) are for homocyclic-based analyses, and the right panels are for total PAC-based analyses ..... 204
**Figure 8.1:** Complete non-targeted data analysis methodology used to identify polycyclic aromatic compounds ..... 241
**Figure 8.2:** The steps used in the aromaticity index (AI) filter in workflow 2 ..... 247
**Figure 8.3:** Kendrick mass defect plots created for halogenated PACs standards using $Br_2$ (top panel), $Cl_2$ (middle panel), and $CH_2$ (bottom panel). Colored bands show 95% predication intervals for each compound class. ..... 257

# LIST OF TABLES

**Table 1.1:** Summary of physicochemical properties of the United States Environmental Protection Agency 16 priority polycyclic aromatic hydrocarbons. ..... 4

**Table 1.2:** Summary of physicochemical properties of some common hetero-polycyclic aromatic compounds ..... 13

**Table 1.3:** The environmental distribution data of hetero-polycyclic aromatic compounds ..... 24

**Table 1.4:** Experimental evidence on toxic effects of hetero-polycyclic aromatic compounds ... 28

**Table 2.1:** The targeted analytes and their multiple reaction monitoring parameters used in this study ..... 62

**Table 2.2:** Characteristics of the four multivariate approaches used in this study ..... 80

**Table 3.1:** Analyte, elution time, product ions, MS/MS quantitation and confirmation ions and collision energies (CE 1 and CE 2) used for HPACs and halo-HPACs. ..... 99

**Table 3.2:** Performance characteristics of the HPACs and halo-HPACs method ..... 107

**Table 3.3:** Mean concentrations (ng/g, wet and dry weight, *n=3*) of HPACs and halo-HPACs in Randle Reef sediments and percentage of analytes observed by sequential extractions using a 50:50 mixture of DCM:ethyl acetate ..... 111

**Table 5.1:** Arithmetic mean (± SD), median and range of the lipid normalized concentrations (µg/g) of different PAC groups detected in the four fish species collected from the four regions of Lake Erie ..... 168

**Table 5.2:** One way ANOVA/ Kruskal-Walli's and Post Hoc test results indicating the region-dependent concentration variations of PAC classes1. ..... 171

**Table 6.1:** Information on the multivariate analyses done for the identification of the species clustering based on the PAC profiles ..... 205

**Table 6.2:** Polycyclic aromatic compounds that showed statistically significant TMF values ($p<0.05$) in the Lake Erie food web. Regression parameters for each analyte are also provided. ..... 214

**Table 6.3:** Three-tier prioritization pollutant framework of PACs for Lake Erie fish ..... 216

**Table 8.1:** Halogen class assignment criteria used in workflow 2 ..... 246

# APPENDIX

**Table I 1:** General information regarding the standards included in the two MRM methods. .....II

**Table I 2:** Percent recoveries of representative HPACs extracted (n=3) using DCM and DCM:ethyl acetate 1:1. ..... VI

**Table I 3:** Mean percent recoveries (± standard deviation, n = 8) of RIS based on IPIS (anthracene-$d_{10}$) during the method validation ..... VII

**Table I 4:** Mean percent recoveries (±SD) of native PAHs with the extraction method developed for HPACs. ..... VIII

**Table I 5:** Calibration curve information of the HPACs and Halo-HPACs standards (pg/µL, n=3) ..... IX

**Table I 6:** Mean concentrations and percent relative standard deviation (% RSD) of analytes in the six procedural blank samples used in the method validation. ..... XI

**Table I 7:** Mean concentrations ± standard deviations (µg/g, dry weight, n=3) of 16 PAHs and 30 APAHs in sediment sample from Hamilton Harbour. ..... XIII

**Table I 8:** Concentrations of analytes (ng/g) in the procedural blanks (n = 3) used in the processing of Randle Reef sediments. .....XV

**Table II 1:** Target analytes and internal standards used for the analysis.....XVIII

**Table II 2:** Sample names and sampling information for sediments used in this study (HH - Hamilton Harbour, TH - Toronto Harbour and NH – Non Harbour)..... XXI

**Table II 3:** Detection percentages calculated from considering all the samples and reference materials.....XXIII

**Table II 4:** Concentrations (ng/g, dry weight) of HPAC in sediments in each year from each sampling site ..... XXIV

**Table II 5:** Mass accuracies in detected analytes calculated from using five random samples .....XXVIII

**Table II 6:** Summary of screening criteria for analytes suitable for chemical fingerprinting XXIX

**Table II 7:** The Spearman's rank correlation rho values, p-values and $R^2$ values for each analyte and HPACs type .....XXXIII

**Table II 8:** Screening criteria for the groupings of reference materials ..... XXXV

**Table III 1:** The targeted analytes and their reference codes used in this study ..... XL
**Table III 2:** The targeted analytes and their CAS numbers ..... XLII
**Table III 3:** Average internal standard recoveries for all data calculated with respect to IPIS (Anthracene-d10) ..... XLIV
**Table III 4:** Average lipid percentage of each species. ..... XLIV
**Table III 5:** Performance characteristics of the extraction and purification method used for HPACs ..... XLVIII
**Table III 6:** Analyte-species combinations that have significantly different concentrations in fish from eastern and western regions ..... L
**Table III 7**: The selected analytes for UMAP and PCA and their loading scores for the first five principal components ..... LVI

**Table IV 1:** The targeted analytes and their reference codes used in this study ..... LIX
**Table IV 2:** The targeted analytes and their CAS numbers ..... LXI
**Table IV 3:** The sampling information and isotope data for the fish samples ..... LXIII
**Table IV 4:** Performance characteristics of the extraction and purification method used for HPACs ..... LXIX
**Table IV 5:** Average lipid percentage of each species ..... LXXI
**Table IV 6:** The lipid normalized concentrations (ng/ g) of each PAC groups in Lake Eire western basin samples expressed as median and mean ± standard deviation ..... LXXII
**Table IV 7:** The calculated biomagnification factors considering major prey-predator relationships ..... LXXIV
**Table IV 8:** The calculated tropic magnification factors considering all species ..... LXXX

**Table V 1:** The targeted analytes, internal standards and their CAS numbers ..... LXXXIV
**Table V 2:** Sample names and sampling information for sediments used in this study (HH - Hamilton Harbour) ..... LXXXVI

**Table V 3**: PACs library created using the authentic standards ..... XCI
**Table V 4:** The results of a standard processed through the workflow 1 .....XCIV
**Table V 5:** Aromaticity indices for the targeted analytes .....XCVIII
**Table V 6:** The results obtained for benchmarking step of aromaticity index filtering ..... CI
**Table V 7:** Possible halogenated PACs that were incorporated in workflow 2 and their exact masses ..... CI
**Table V 8:** Summary of the output file of the standard run checked through the workflow 2.. CXI
**Table V 9:** Processed output from the KMD assessment of the standard data obtained at workflow 2.....CXVI
**Table V 10:** HPACs reference masses used in workflow 3 .....CXIX
**Table V 11**: The performance characteristics of WF3 at different evidence score thresholds .....CXXIII
**Table V 12**: A comparison of the features and performance characteristics between W2 and WF3 .....CXXIII
**Table V 13:** Compounds detected in the environmental samples .....CXXIV
**Table V 14:** Concentrations of non-targeted analytes calculated using the relative response of neighboring target analytes in sediment samples.....CXXVI

**Figure II 1:** Geographical locations of the sampling sites created using Google My Maps .....XX
**Figure II 2:** Concentration variances of Hamilton Harbour suspended sediment samples from 2014 to 2023 ..... XXVI
**Figure II 3:** The concentration differences of HPACs in suspended sediments from HH site 9031 (a) with indole (top panel) and (b) without indole (bottom panel) .....XXVII
**Figure II 4:** DBT ratios calculated for the reference samples (CT RM - blue, NCT RM -orange and Mixed RM - green)..... XXXI
**Figure II 5:** Ratios between dibenzothiophene and its alkylated derivatives in HH suspended sediments.....XXXII
**Figure II 6:** Scatter plots between the total combustion PAHs concentration and concentrations of (a) PANHs (top panel), (b) PASHs (middle panel) and (c) PAOHs (bottom panel) in each sediment sample. ..... XXXIV

**Figure II 7:** Biplot for the principal component analysis ..... XXXVI
**Figure II 8:** PCA and UMAP analysis outputs between (a) PC1 and PC2 (top panel), (b) PC1 and PC3 (bottom panel) for HH harbour suspended sediments colored based on different sampling sites, the CT reference materials (yellow dots), the NCT reference materials (orange dots) and the Mixed RM (dark blue dots)..... XXXVII
**Figure II 9:** PCA Loading values for each analyte .....XXXVIII

**Figure III 1**: Sample locations from Lake Erie..... XLVI
**Figure III 2**: PANH concentrations detected in each species (FD: freshwater drum, GS; gizzard shad, YP: yellow perch, WE: walleye, WE: western, WC: west central, EC: east central, ES: eastern)..... LII
**Figure III 3**: PASH and PAOH concentrations detected in each species (FD: freshwater drum, GS; gizzard shad, YP: yellow perch, WE: walleye, WE: western, WC: west central, EC: east central, ES: eastern). .....LIII
**Figure III 4**: Lipid corrected PAH concentrations detected in each species (FD: freshwater drum, GS; gizzard shad, YP: yellow perch, WE: walleye, WE: western, WC: west central, EC: east central, ES: east.....LIV
**Figure III 5**: Lipid corrected APAH concentrations detected in each species (FD: freshwater drum, GS; gizzard shad, YP: yellow perch, WE: walleye, WE: western, WC: west central, EC: east central, ES: eastern). ..... LV
**Figure III 6**: Heatmaps generated from hierarchical cluster analysis results of all four species collected from the four regions of Lake Erie .....LVII

**Figure IV 1**: The UMAP analyses output graphs generated considering variable 1 and variable 3 with (a) only HPACs, (b) with PAHs and APAHs and (c) all PACs ..... LXXIII
**Figure IV 2:** Plots between the $\log_{10}$concentation and tropic positions that give statistically significant biomagnifying PACs .....LXXIX

**Figure V 1:** Kendrick mass defect plots created for halogenated PACs standards using H/Cl (top panel), $Br_2$ (middle panel), and $CH_2$ (bottom panel). Colored bands show 95% predication intervals for each compound class. ..... CXIII
**Figure V 2:** Kendrick mass defect plots generated for the two sediment samples ..... CXIV
**Figure V 3:** Kendrick mass defect plots generated for the two fish samples ..... CXV

# LIST OF ABBREVIATIONS

| Abbreviations | Definition |
|---|---|
| ACF | Auto correlation factors |
| AED | Atomic emission detector |
| AI | Aromaticity Index |
| AhR | Aryl hydrocarbon receptor |
| ANOVA | Analysis of Variance |
| AOSR | Athabasca oil sands region |
| APAHs | Alkylated polycyclic aromatic hydrocarbons |
| ARF | Average response factor |
| ASE | Accelerated solvent extraction |
| BCF | Bioconcentration factor |
| BMF | Biomagnification factor |
| BP | Boiling point |
| CE | Collision energy |
| CI | Chemical ionization |
| CI | Confidence interval |
| CID | Collision-induced dissociation |
| CT | Coal-tar |
| CRMs | Certified reference materials |
| CSOs | Combined sewer outfalls |
| CT RM | Coal-tar based reference materials |
| DAD | Diode array detector |
| DBTs | Dibenzothiophenes |
| DC | Direct current |
| DCM | Dichloromethane |
| DE | Diatomaceous earth |
| DE | Double bond equivalence |
| DM | Dreissenid mussels |
| ECNI | Electron capture negative ionization |
| EI | Electron ionization |
| ER | Estrogen receptor |
| ES | Emerald shiner |
| ESI | Electrospray ionization |
| FD | Freshwater drum |
| FID | Flame ionization detection |
| FS | Full scan |
| FT-ICR | Fourier transform ion cyclotron resonance |

| | |
|---|---|
| FWHM | Full width at half maximum |
| GC×GC–TOF-MS | Two-dimensional gas chromatography with time-of-flight mass spectrometry |
| GC-MS | Gas chromatography mass spectrometry |
| GC-MS/MS | Gas chromatography coupled with tandem mass spectrometry |
| GDP | Gross domestic product |
| GLIER | Great Lakes Institute for Environmental Research |
| GPC | Gel permeation chromatography |
| GS | Gizzard shad |
| HCA | Hierarchical cluster analysis |
| HH | Hamilton harbour |
| HMW | High molecular weight |
| HPACs | Hetero-polycyclic aromatic compounds |
| HPAHs | Halogenated polycyclic aromatic hydrocarbons |
| HPLC | High performance liquid chromatography |
| hr | Hours |
| HRMS | High-resolution mass spectrometry |
| HSD | Honestly significant difference |
| IARC | International Agency for Research on Cancer |
| IPIS | Instrument performance internal standard |
| $K_a$ | Acid dissociation constant |
| $K_{oa}$ | Octanol-air partition coefficient |
| $K_{oc}$ | Soil adsorption coefficient |
| $K_{ow}$ | Octanol-water partition coefficient |
| LMW | Low molecular weight |
| LOD | Lower detection limit |
| lw | Lipid corrected weight |
| m/z | Mass to charge ratio |
| MP | Melting point |
| MRM | Multiple reaction monitoring |
| NCT RM | Non-coal-tar based reference materials |
| NH | Non-harbour |
| NCT | Non-coal-tar |
| NTA | Non-targeted analysis |
| OLSR | Ordinary least squares regression |
| PACs | Polycyclic aromatic compounds |
| PAHs | Polycyclic aromatic hydrocarbons |
| PANHs | Polycyclic aromatic nitrogen heterocycles |
| PAOHs | Polycyclic aromatic oxygen heterocycles |
| PASHs | Polycyclic aromatic sulfur heterocycles |

| | |
|---|---|
| PBDEs | Polybrominated diphenyl ethers |
| PBT | Persistence, bioaccumulation and toxicity |
| PC1 | Principal component 1 |
| PC2 | Principal component 2 |
| PC3 | Principal component 3 |
| PCA | Principal component analysis |
| PCBs | Polychlorinated biphenyls |
| p-chem | Physicochemical |
| PFTBA | Perfluorotributylamine |
| PIS | Production ion scan |
| PIS | Product ion scan |
| PLE | Pressurized solvent extraction |
| POP | Persistent organic pollutants |
| ppm | Parts per million |
| Q1 | First quadrupole |
| q2 | Collision cell |
| Q3 | Third quadrupole |
| QA/QC | Quality assurance and quality control |
| RDA | Redundancy analysis |
| RF | Radio frequency |
| RF and rf | Response factor |
| RG | Round goby |
| RIS | Recovery internal standard |
| ROS | Reactive oxygen species |
| RSD | Relative standard deviation |
| rt | Retention time |
| rt1 | First dimension retention time |
| rt2 | Second dimension retention time |
| S/N | Signal-to-noise |
| SD | Standard deviation |
| SIM | Selected ion monitoring |
| SS | Spottail shiner |
| $t_{1/2}$ | Half-life |
| TH | Toronto harbour |
| TMF | Tropic magnification factor |
| TMF | Time of flight |
| TP | Tropic position |
| UHP | Ultra-high purity |
| UMAP | Uniform manifold approximation and projection |
| US EPA | United states environmental protection agency |

| | |
|---|---|
| WB | White bass |
| WE | Walleye |
| WF1 | Workflow 1 |
| WF2 | Workflow 2 |
| WF3 | Workflow 3 |
| WS | White sucker |
| WWTP | Wastewater treatment plant |
| YP | Yellow perch |

# CHAPTER 1 INTRODUCTION

## 1.1 POLYCYCLIC AROMATIC COMPOUNDS

The oil and gas industry is an integral component of the Canadian economy, contributing substantially to national gross domestic product (GDP) and energy production. Statistics Canada and federal energy reports indicate that the sector contributes approximately 4–5% of Canada's GDP and that crude oil production was around 5 million barrels per day in 2023.[1,2] Industrial activity on this scale inevitably results in the release of oil into Canadian environmental systems through processes such as natural seepage, accidental oil spills, and operational discharges.[3]

Crude oil is a complex mixture of thousands of chemicals. Generally, crude oil contains four major classes of compounds known by the acronym SARA: saturates, aromatics, resins, and asphaltenes, as well as minor constituents such as organometallic compounds (e.g., petroporphyrins), naphthenic acids, heavy metals, and some minerals. When the elemental composition is considered, crude oil mainly contains aliphatic and aromatic hydrocarbons and compounds containing heteroatoms such as nitrogen, sulfur, and oxygen, in lesser amounts.[3]

Polycyclic aromatic compounds (PACs) are a major class of compounds found in crude oil, consisting of one or more fused aromatic rings primarily containing carbon and hydrogen. Among the numerous classes of organic environmental pollutants, PACs have remained a focus of environmental scientists for several decades.[4] In addition to petrogenic sources, such as crude oil and other petroleum products, PACs also originate from pyrogenic and biogenic processes. Pyrogenic processes are involved in the incomplete combustion of organic matter during events

such as wildfires, vehicle exhaust, and volcanic eruptions. Biogenic processes contribute relatively less to the overall environmental burden of PACs compared to the other two processes.[5]

Scientific interest in these compounds arises from their natural occurrence, considerable structural diversity, which leads to a wide range of physicochemical (p-chem) properties, their chemical complexity, and their known persistence, bioaccumulation potential, and toxic (PBT) effects.[6] Based on the structural characteristics, PACs can be divided into several subclasses, namely polycyclic aromatic hydrocarbons (PAHs), alkylated polycyclic aromatic hydrocarbons (APAHs), halogenated polycyclic aromatic hydrocarbons (HPAHs), and hetero-polycyclic aromatic compounds (HPACs).[7]

Polycyclic aromatic hydrocarbons are characterized by fused aromatic ring structures composed exclusively of carbon and hydrogen atoms. Several decades ago, with the discovery of toxic effects and widespread occurrence of several of these compounds, scientists highlighted the importance of regulating PAHs. As a result, in 1978, the United States Environmental Protection Agency (US EPA) identified 16 of these PAHs (Figure 1.1) as priority pollutants.[8] Subsequently, these chemicals were included in the Canadian Environmental Protection Act, 1999, as part of the Toxic Substances List.[9] Table 1.1, section 1, summarizes the p-chem properties and other PBT-related characteristics of the 16 PAHs.

As their names suggest, APAHs and HPAHs contain alkyl and halogen substituents, respectively, attached to the PAH backbone. The structural complexity and diversity of these subclasses are much greater than those of PAHs, resulting in a substantial gap in our understanding of these compounds compared to traditional PAHs.[10] Numerous studies have shown that APAHs are more abundant than PAHs in most environmental matrices.[11] Moreover, several APAHs and HPAHs

have been found to exhibit higher toxicity than PAHs to humans and other organisms.[12] Despite these findings, the original 16 PAHs remain the only priority pollutants among all PACs. [13]

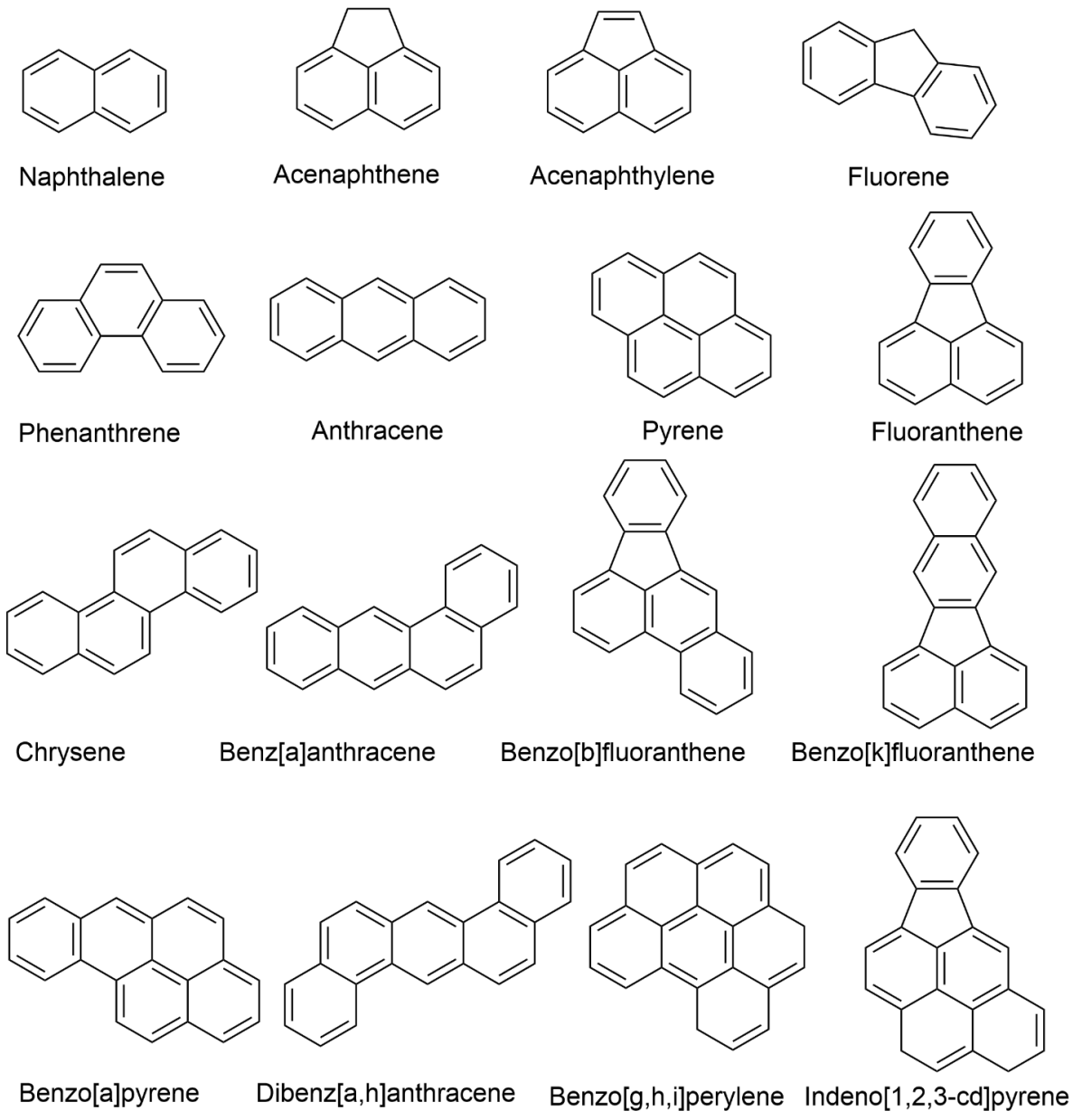



**Figure 1.1:** The chemical structures of the United States Environmental Protection Agency's 16 priority polycyclic aromatic hydrocarbons.

**Table 1.1:** Summary of physicochemical properties of the United States Environmental Protection Agency 16 priority polycyclic aromatic hydrocarbons.

| Compound | CAS number | log $K_{ow}$[a] | MP[b] (°C) | BP[c] (°C) | Water Solubility (mg/L) | Vapor pressure (mmHg) | Henry's Law Constant (atm.m³/mol) | Log $K_{oa}$[d] | Log $K_{oc}$[e](L/kg) | Half-life in water (hr) | Half-life in sediments (hr) | Log BCF[f] | log BMF[g] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Naphthalene | 91-20-3 | 3.17 | 5 | 232 | 184 | $7 \times 10^{-2}$ | $3.7 \times 10^{-4}$ | 4.99 | 3.19 | 900 | $8.1 \times 10^{-3}$ | 1.76 | 2.16 |
| Acenapthene | 83-32-9 | 3.92 | 93 | 279 | 3.9 | $2 \times 10^{-3}$ | $1.8 \times 10^{-4}$ | 6.31 | 3.59 | 900 | $8.1 \times 10^{-3}$ | 2.25 | 1.98 |
| Acenapthylene | 208-96-8 | 3.94 | 92 | 280 | 16.1 | $7 \times 10^{-3}$ | $1.1 \times 10^{-4}$ | 6.27 | 3.75 | 360 | $3.2 \times 10^{-3}$ | 2.27 | 2.76 |
| Fluorene | 86-73-7 | 4.18 | 115 | 295 | 1.69 | $6 \times 10^{-4}$ | $9.6 \times 10^{-5}$ | 6.79 | 3.70 | 360 | $3.2 \times 10^{-3}$ | 2.42 | 2.63 |
| Phenanthrene | 85-01-8 | 4.46 | 99 | 340 | 1.15 | $1 \times 10^{-4}$ | $4.2 \times 10^{-5}$ | 7.57 | 4.35 | 1440 | $1.3 \times 10^{-4}$ | 3.27 | 2.90 |
| Anthracene | 120-12-7 | 4.54 | 216 | 342 | 0.076 | $7 \times 10^{-6}$ | $2.5 \times 10^{-5}$ | 7.78 | 4.40 | 1440 | $1.3 \times 10^{-4}$ | 3.30 | 2.88 |
| Fluoranthene | 206-44-0 | 5.22 | 111 | 375 | 0.26 | $9 \times 10^{-6}$ | $6.5 \times 10^{-6}$ | 8.80 | 4.95 | 5760 | $5.2 \times 10^{-4}$ | 3.68 | 2.35 |
| Pyrene | 129-00-0 | 5.18 | 150 | 393 | 0.132 | $5 \times 10^{-6}$ | $5.2 \times 10^{-6}$ | 8.87 | 4.88 | 5760 | $5.2 \times 10^{-4}$ | 3.66 | 2.38 |
| Benz[a]anthracene | 56-55-3 | 5.91 | 158 | 438 | 0.011 | $6 \times 10^{-7}$ | $4.8 \times 10^{-6}$ | 9.65 | 5.52 | 8760 | $7.9 \times 10^{-4}$ | 4.05 | 2.14 |
| Chrysene | 218-01-9 | 5.86 | 254 | 448 | 0.0015 | $8 \times 10^{-9}$ | $5.5 \times 10^{-6}$ | 9.57 | 5.48 | 8760 | $7.9 \times 10^{-4}$ | 4.02 | 2.15 |
| Benzo[a]pyrene | 50-32-8 | 6.04 | 179 | 496 | 0.0038 | $7 \times 10^{-7}$ | $4.6 \times 10^{-7}$ | 10.4 | 6.11 | 17520 | $1.6 \times 10^{-5}$ | 4.15 | 2.09 |
| Benzo[b]fluoranthene | 205-99-2 | 5.8 | 168 | 481 | 0.0012 | $7 \times 10^{-7}$ | $6.2 \times 10^{-7}$ | 9.94 | 5.80 | 17520 | $1.6 \times 10^{-5}$ | 4.09 | 2.12 |
| Benzo[k]fluoranthene | 207-08-9 | 6 | 216 | 480 | 0.0007 | $1 \times 10^{-10}$ | $4.0 \times 10^{-8}$ | 10.4 | 6.05 | 17520 | $1.6 \times 10^{-5}$ | 4.14 | 2.10 |
| Indeno[1,2,3-cd]pyrene | 193-39-5 | 6.58 | 163 | 536 | 0.062 | $2 \times 10^{9}$ | $2.0 \times 10^{-8}$ | 11.1 | 6.37 | 17520 | $1.6 \times 10^{-5}$ | 4.25 | 2.06 |
| Dibenzo[a,h]anthracene | 53-70-3 | 6.84 | 270 | 524 | 0.0005 | $1 \times 10^{10}$ | $1.6 \times 10^{-8}$ | 11.4 | 6.52 | 17520 | $1.6 \times 10^{-5}$ | 4.30 | 2.05 |
| Benzo[g,h,i]perylene | 191-24-2 | 6.63 | 277 | 500 | 0.0003 | $1 \times 10^{10}$ | $1.5 \times 10^{-8}$ | 11.2 | 6.42 | 17520 | $1.6 \times 10^{-5}$ | 4.27 | 2.06 |

a – octanol-water partition coefficient, b – melting point, c – boiling point, d – octanol-air partition coefficient, e – soil adsorption coefficient, f – bioconcentration factor, and g – biomagnification factor (The values were obtained from US EPA ECOSAR database https://episuite.dev/#/)

## 1.2 HETERO-POLYCYCLIC AROMATIC COMPOUNDS

According to Balaban *et al.*,[14] at the beginning of the 21st century, *ca*. 20 million chemicals were known, of which more than two-thirds were fully or partially aromatic, and about half were heteroaromatic. Hetero-polycyclic aromatic compounds are a subclass of PACs that have at least one nitrogen, sulfur, and/or oxygen atom replacing one or more carbon atoms in the aromatic structure. These compounds occur naturally in petroleum as minor constituents (e.g., 0.1-2% nitrogen, 0.05-6% sulfur) alongside other PAC classes.[15,16]

Hetero-polycyclic aromatic compounds are generally classified based on the heteroatom as polycyclic aromatic nitrogen heterocycles (PANHs), sulfur heterocycles (PASHs), and oxygen heterocycles (PAOHs). Common PANHs include quinolines, carbazoles, and acridines; typical PASHs include benzothiophenes and dibenzothiophenes; while benzofurans and dibenzofurans are the common representatives of PAOHs.[17] Some common examples of unsubstituted HPACs are shown in Figure 1.2.

Depending on the heteroatom type, the structural diversity of alkylation and halogenation in HPACs is much wider than in other PAC subclasses. Unfortunately, this also results in a lack of scientific studies pertaining to these compounds. Key challenges include the limited availability of commercial standards, the vast number of possible structural isomers, the lack of certified reference materials (CRMs), and associated economic constraints.[18] Also, the analytical challenges of HPACs analysis arise from the coelution and their spectral similarities with PAHs and APAHs. Over the last few decades, the improved detection capabilities of advanced analytical techniques such as triple-quadrupole tandem mass spectrometry, high-resolution mass spectrometry, and

multidimensional gas chromatography have enabled more accurate analysis of HPACs in environmental samples.[4,18,19]

Due to these analytical challenges, studies aimed at evaluating the PBT characteristics of HPACs are also limited. However, available evidence suggests that HPACs often exhibit PBT properties comparable to or even greater than those of parent PAHs.[12] The incorporation of heteroatoms alters key p-chem properties, such as polarity, solubility, and reactivity, of HPACs, resulting in distinct environmental fate and distribution compared to other classes of PACs.[20,21]

Because of the above reasons, research focused on HPACs is essential to fully understand the overall environmental impact of PACs. This chapter introduces the fundamental characteristics and sources of HPACs, summarizes the current state of their environmental occurrence and toxicity, and outlines the scope of my thesis.

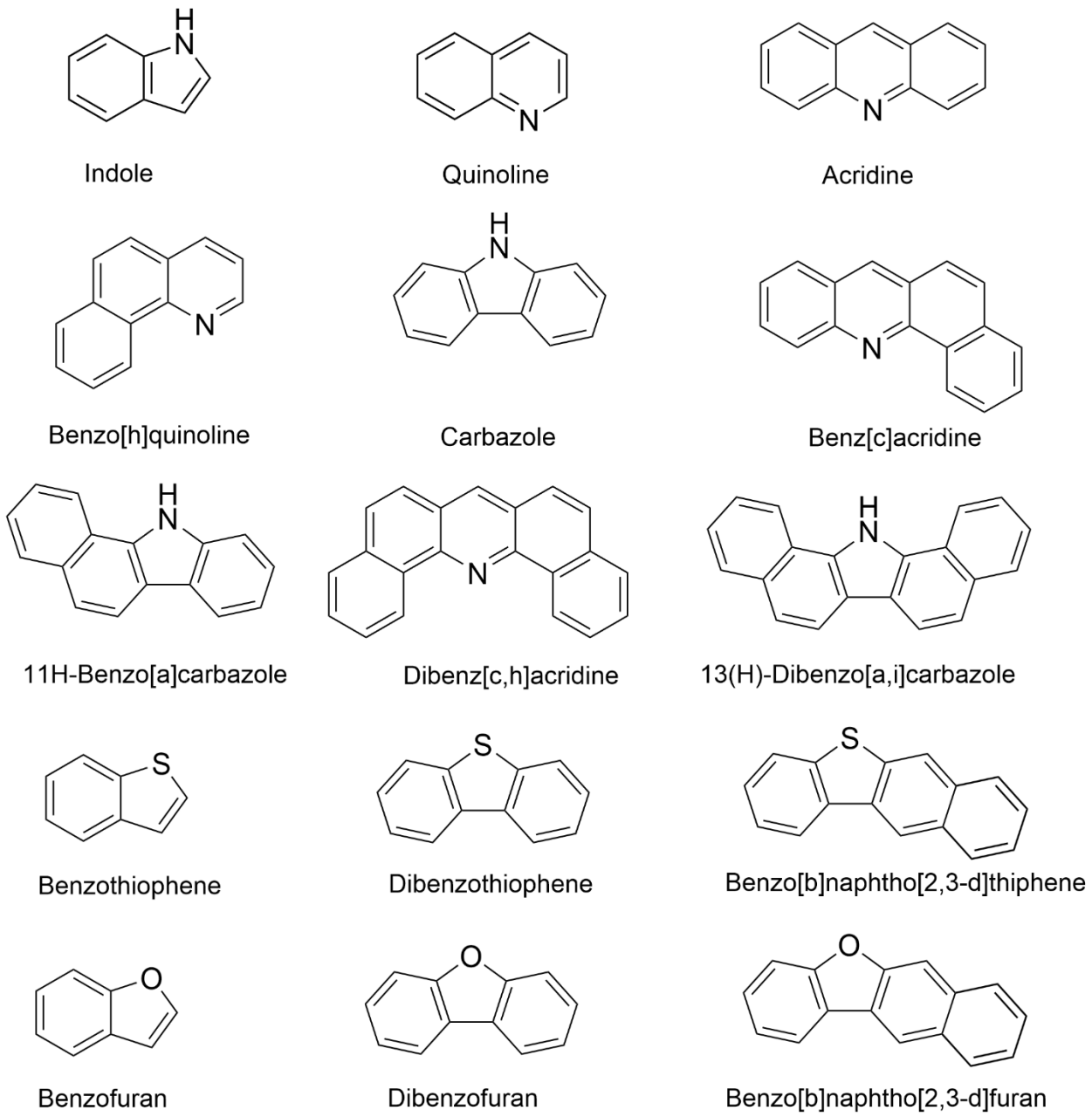



**Figure 1.2:** The chemical structures of some common unsubstituted hetero-polycyclic aromatic compounds

## 1.3 ELECTRONIC AND PHYSICOCHEMICAL PROPERTIES OF HETERO-POLYCYCLIC AROMATIC COMPOUNDS

Aromaticity is a concept that was first introduced in the mid-19th century to distinguish benzene from unsaturated hydrocarbons.[14] The key traits of aromatic systems include low reactivity arising from thermodynamic stability via resonance energy, high magnetic susceptibility, bond-length equalization, and π-electron delocalization, yielding stability, equalized bonds, ring currents, and a preference for electrophilic substitution.[22,23]

Polycyclic aromatic compounds extend the concept of aromaticity to compounds containing two or more fused aromatic rings. This results in larger, planar, conjugated π-electron systems that are more stable than one-ring aromatic systems such as benzene, pyridine, pyrrole, thiophene, and furan.[24–27] With increasing molecular size, PACs generally exhibit decreased volatility and aqueous solubility (i.e., increased hydrophobicity), leading to strong partitioning into organic phases, such as sediments and lipid-rich animal tissues.[4] These p-chem properties govern the persistence and bioaccumulation characteristics of these molecules. Additionally, the distribution of electron density across the fused-ring system governs their chemical reactivity, including regioselectivity in electrophilic substitution and susceptibility to oxidative transformations. [28–30]

The substitution of a carbon with a heteroatom such as nitrogen, sulfur, or oxygen notably modifies the electronic characteristics of PACs. This leads to changes in p-chem properties of these compounds compared to their hydrocarbon counterparts. Heteroatoms introduce an electronegativity difference into the ring system, thereby altering its π electron distribution, molecular polarity, and chemical reactivity.[31,32] The electronic properties of HPACs vary primarily with the heteroatom.

Environmentally relevant nitrogen-containing HPACs mainly have either a six-membered pyridine ring or a five-membered pyrrole ring in their polycyclic aromatic structures. Depending on the number of atoms in these rings, the electronic effect of the nitrogen atom substantially changes. In both cases, nitrogen is $sp^2$ hybridized.

In pyridine-containing PANHs, the unhybridized p-orbital contains a single electron, which contributes to the delocalized π-electron system. Conversely, a lone pair of electrons on N occupies an $sp^2$ orbital, which is not part of the aromatic system. Nitrogen is also a highly electronegative element, therefore the π-electron density in the ring becomes reduced. Hence, the reactivity towards electrophilic substitution decreases compared to pyrrole-containing PANHs. The presence of a lone pair on the nitrogen atom makes pyridine a weak base. Also, the increased polarity from these electronic effects leads to greater water solubility and bioavailability of these compounds relative to their homocyclic analogues.[14,28]

In pyrrole, nitrogen participates in an N-H bond, and therefore, the lone pair is in a p-orbital, which is perpendicular to the ring. Hence, the lone pair contributes to the aromatic system, increasing the electron density and electrophilic substitution reactivity of the ring. However, the unavailability of the lone electron pair makes pyrrole-containing PANHs less basic and less polar than pyridine-based PANHs.[33,34]

In the environment, sulfur and oxygen-containing HPACs exist as five-membered thiophenes and furans, respectively. Since both sulfur and oxygen have a similar number of valence electrons, the basic electronic structures of these compounds are similar. Both sulfur and oxygen are $sp^2$ hybridized and have two lone pairs of electrons - one in the p-orbital contributing to the aromatic system and one in an unhybridized $sp^2$ orbital. However, oxygen is a highly electronegative

element and has a greater 2p–2p orbital overlap with carbon atoms in the ring than sulfur (3p–2p overlap). These factors make furans more reactive toward electrophilic substitution and more polar than thiophenes. The higher electronegativity of oxygen, however, also means that its lone pair is held more tightly, making furan slightly less basic than thiophene despite having an in-plane $sp^2$ lone pair.[14,28,35]

The effect of heteroatom substitution on HPACs is governed by the extent to which heteroatom lone pairs of electrons participate in the aromatic π-system and the efficiency of orbital overlap.[36] These factors together control the electron density, reactivity, basicity, and ultimately the environmental behavior of these compounds.[12]

The replacement of a heteroatom and the electronic changes introduced thereby directly influence the p-chem properties that govern the environmental fate and behavior of HPACs. Some of these parameters include aqueous solubility, vapor pressure, octanol-water partition coefficient ($K_{ow}$), organic carbon-water partition/sorption coefficient ($K_{oc}$), and acid dissociation constant ($K_a$).[37]

Pyridine-containing PANHs, such as quinolines and acridines, tend to have relatively high polarity and water solubility due to the presence of a non-aromatic lone pair and electron-withdrawing effects.[38] This facilitates their partitioning into the aqueous phase, increasing the environmental mobility and bioavailability.[39] Meanwhile, pyrrole-containing PANHs such as indole and carbazole are generally less polar and more hydrophobic, resulting in stronger partitioning into organic matter and lipid phases than their pyridine-containing counterparts.[40,41]

Polycyclic aromatic sulfur heterocycles (e.g., thiophenes and dibenzothiophenes) typically exhibit p-chem properties like PAHs due to their lower polarity and weaker electron-donating character. However, the higher polarizability of sulfur enhances intermolecular interactions, often leading to

strong sorption to sediments and organic matter.[42,43] Furans (PAOHs) generally exhibit higher polarity and reactivity due to strong resonance stabilization and electronegative effects, resulting in increased aqueous solubility and susceptibility to photochemical and oxidative transformations.[44–46]

Because the environmental fate and behavior of HPACs changes with these p-chem properties, properties such as environmental half-life ($t_{1/2}$), bioaccumulation factor (BAF), biomagnification factor (BMF), toxicity, biodegradability, and long-range transport characteristics of HPACs can vary among PAC classes.[47–49] For instance, compounds with high $K_{ow}$ values tend to accumulate in sediments and biota, thereby increasing their $t_{1/2}$'s and bioaccumulation. More polar HPACs are more mobile in aquatic systems and may exhibit greater bioavailability and a higher potential for transport in aqueous environments. Additionally, the electronic structure influences the degradation pathways of these compounds. Electron-rich compounds are more susceptible to electrophilic and oxidative transformations; meanwhile, electron-deficient systems may exhibit greater persistence under certain environmental conditions.

Compared with their parent HPACs, alkylated and halogenated derivatives exhibit distinct p-chem properties that depend on the nature and extent of substitution. Generally, alkyl groups increase the hydrophobicity by adding nonpolar carbon groups and enhancing dispersion interactions.[46,50,51] Therefore, alkylation typically results in greater $K_{ow}$, Henry's constant, $K_{oc}$, BCF, and BMF values, and lower biodegradability relative to their non-alkylated analogues.[52]

Halogenation also tends to increase the hydrophobicity and molecular size of PACs, resulting in greater $K_{ow}$ values, stronger sorption to organic matter, and enhanced bioaccumulation potential.[53] However, the effect of halogen substitution on polarity and electron distribution is more complex.

While halogens exert a strong electron-withdrawing inductive effect, they can also participate in resonance depending on the system.[54] In HPACs, halogenation may alter electron density distribution across both the aromatic ring and heteroatom, influencing reactivity and degradation pathways. Overall, halogenated HPACs are often more persistent due to increased chemical stability and resistance to biodegradation.[55,56]

Consequently, both alkylation and halogenation generally promote environmental persistence, bioaccumulation, and, in many cases, toxicity, although the magnitude of these effects depends on molecular structure, substitution pattern, metabolic processes and environmental conditions.

Overall, the interplay between electronic structure and p-chem properties determines the environmental fate of HPACs, including their transport, partitioning, transformation, and biological interactions. Understanding these relationships is essential for accurately assessing the environmental risk and behavior of HPACs in complex environmental matrices. Table 1.2 compares the main p-chem properties and PBT characteristics of commonly found HPACs and their alkylated and halogenated derivatives.

**Table 1.2:** Summary of physicochemical properties of some common hetero-polycyclic aromatic compounds

| Compounds Name | CAS number | log $K_{ow}$[a] | MP[b] (°C) | BP[c] (°C) | Water Solubility (mg/L) | Vapor pressure (mmHg) | Henry's Law Constant (atm.m³/mol) | Log $K_{oa}$[d] | Log $K_{oc}$[e] (L/kg) | Half-life in water (hr) | Half-life in sediments (hr) | Log BCF[f] | log BMF[g] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quinoline | 91-22-5 | 2.03 | -15 | 237.6 | 6110 | $9.1 \times 10^{-3}$ | $2.49 \times 10^{-3}$ | 5.64 | 2.2 | 900 | $8.1 \times 10^{3}$ | 0.97 | 1.42 |
| 2-Methylquinoline | 91-63-4 | 2.59 | -2 | 246.5 | 620 | $3.5 \times 10^{-3}$ | $3.10 \times 10^{-7}$ | 6.08 | 2.52 | 900 | $8.1 \times 10^{3}$ | 1.34 | 1.60 |
| 2-Chloroquinoline | 612-62-4 | 3.12 | 37 | 266 | 61 | $6.8 \times 10^{-4}$ | $5.20 \times 10^{-7}$ | 6.86 | 3.07 | 1440 | $1.3 \times 10^{4}$ | 1.79 | 1.90 |
| 2-Bromoquinoline | 2437-83-4 | 3.43 | 24 | 278 | 28 | $3.2 \times 10^{-4}$ | $5.80 \times 10^{-7}$ | 7.23 | 3.33 | 1440 | $1.3 \times 10^{4}$ | 2.01 | 2.00 |
| Isoquinoline | 119-65-3 | 2.08 | 27 | 243.2 | 4520 | $7.5 \times 10^{-3}$ | $2.58 \times 10^{-7}$ | 5.72 | 2.26 | 900 | $8.1 \times 10^{3}$ | 1.00 | 1.44 |
| 3-Methylisoquinoline | 1125-80-0 | 2.63 | 67 | 255 | 510 | $2.8 \times 10^{-3}$ | $3.20 \times 10^{-7}$ | 6.14 | 2.57 | 900 | $8.1 \times 10^{3}$ | 1.38 | 1.62 |
| 3-Bromoisoquinoline | 1532-97-4 | 3.50 | 53 | 281 | 24 | $2.6 \times 10^{-4}$ | $6.00 \times 10^{-7}$ | 7.32 | 3.4 | 1440 | $1.3 \times 10^{4}$ | 2.05 | 2.02 |
| Indole | 120-72-9 | 2.14 | 53 | 253.7 | 3760 | $1.0 \times 10^{-2}$ | $1.60 \times 10^{-7}$ | 5.88 | 2.43 | 360 | $3.2 \times 10^{3}$ | 1.04 | 1.46 |
| 2,3-Dimethylindole | 91-55-4 | 3.09 | 109 | 285 | 232 | $8.0 \times 10^{-4}$ | $2.40 \times 10^{-7}$ | 7.13 | 3.06 | 360 | $3.2 \times 10^{3}$ | 1.72 | 1.84 |
| 5-Chloroindole | 17422-32-1 | 3.01 | 62 | 271 | 190 | $6.5 \times 10^{-4}$ | $3.10 \times 10^{-7}$ | 6.95 | 2.98 | 360 | $3.2 \times 10^{3}$ | 1.66 | 1.81 |
| 5-Bromoindole | 10075-50-0 | 3.34 | 71 | 281 | 96 | $3.2 \times 10^{-4}$ | $3.30 \times 10^{-7}$ | 7.38 | 3.29 | 1440 | $1.3 \times 10^{4}$ | 1.94 | 1.95 |
| Benzo[h]quinoline | 230-27-3 | 3.34 | 52 | 338 | 86 | $2.8 \times 10^{-5}$ | $1.50 \times 10^{-7}$ | 7.65 | 3.48 | 1440 | $1.3 \times 10^{4}$ | 1.97 | 2.01 |
| Acridine | 260-94-6 | 3.40 | 110 | 346 | 38.4 | $8.2 \times 10^{-6}$ | $1.20 \times 10^{-7}$ | 7.88 | 3.37 | 1440 | $1.3 \times 10^{4}$ | 2.06 | 2.04 |
| 2-Methylacridine | 1722-12-9 | 3.95 | 90 | 357 | 9.5 | $3.2 \times 10^{-6}$ | $1.80 \times 10^{-7}$ | 8.42 | 3.87 | 1440 | $1.3 \times 10^{4}$ | 2.36 | 2.12 |
| 9-Chloroacridine | 1207-69-8 | 3.89 | 105 | 367 | 5.8 | $1.8 \times 10^{-6}$ | $2.10 \times 10^{-7}$ | 8.48 | 3.78 | 1440 | $1.3 \times 10^{4}$ | 2.32 | 2.11 |
| 9-bromoacridine | 2150-08-5 | 4.10 | 112 | 376 | 3.1 | $9.5 \times 10^{-7}$ | $2.30 \times 10^{-7}$ | 8.71 | 3.96 | 1440 | $1.3 \times 10^{4}$ | 2.43 | 2.14 |
| 3-Methylbenzo[f]quinoline | 2586-59-6 | 3.88 | 70 | 355 | 10.2 | $3.0 \times 10^{-6}$ | $1.70 \times 10^{-7}$ | 8.40 | 3.77 | 1440 | $1.3 \times 10^{4}$ | 2.31 | 2.10 |
| Carbazole | 86-74-8 | 3.72 | 246 | 354.7 | 1.8 | $1.5 \times 10^{-5}$ | $2.40 \times 10^{-6}$ | 8.12 | 3.96 | 1440 | $1.3 \times 10^{4}$ | 2.26 | 2.56 |
| 9-Methylcarbazole | 1484-12-4 | 4.26 | 189 | 364 | 0.47 | $4.8 \times 10^{-6}$ | $2.90 \times 10^{-6}$ | 8.66 | 4.42 | 5760 | $5.2 \times 10^{4}$ | 2.57 | 2.43 |
| 3,6-Dimethylcarbazole | 6628-93-9 | 4.80 | 158 | 374 | 0.14 | $1.5 \times 10^{-6}$ | $3.50 \times 10^{-6}$ | 9.21 | 4.93 | 5760 | $5.2 \times 10^{4}$ | 2.87 | 2.31 |
| 3,6-Dichlorocarbazole | 6974-16-9 | 4.92 | 210 | 386 | 0.085 | $5.2 \times 10^{-7}$ | $4.10 \times 10^{-6}$ | 9.50 | 5.07 | 5760 | $5.2 \times 10^{4}$ | 2.94 | 2.28 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3,6-Dibromocarbazole | 6825-22-5 | 5.50 | 222 | 402 | 0.022 | $1.4 \times 10^{-6}$ | $4.60 \times 10^{-6}$ | 10.1 | 5.6 | 8760 | $7.9 \times 10^{4}$ | 3.22 | 2.17 |
| 1,2,3,4-Tetrahydrocarbazole | 942-01-8 | 3.85 | 118 | 328 | 3.2 | $3.2 \times 10^{-5}$ | $5.40 \times 10^{-6}$ | 7.45 | 3.78 | 1440 | $1.3 \times 10^{4}$ | 2.30 | 2.10 |
| 3-Tert-butyl-9H-carbazole | 42484-82-2 | 5.06 | 148 | 396 | 0.062 | $6.5 \times 10^{-7}$ | $3.80 \times 10^{-6}$ | 9.52 | 5.1 | 5760 | $5.2 \times 10^{4}$ | 3.00 | 2.26 |
| 9-Phenylcarbazole | 1150-62-5 | 5.42 | 244 | 420 | 0.026 | $1.2 \times 10^{-7}$ | $2.60 \times 10^{-6}$ | 10.2 | 5.48 | 8760 | $7.9 \times 10^{4}$ | 3.18 | 2.18 |
| Benz[c]acridine | 225-51-4 | 4.45 | 96 | 401 | 0.98 | $3.2 \times 10^{-7}$ | $1.80 \times 10^{-7}$ | 9.62 | 4.68 | 5760 | $5.2 \times 10^{4}$ | 2.74 | 2.33 |
| 11H-Benzo[a]carbazole | 239-01-0 | 4.72 | 182 | 411 | 0.3 | $8.5 \times 10^{-8}$ | $1.60 \times 10^{-6}$ | 9.94 | 5 | 5760 | $5.2 \times 10^{4}$ | 2.87 | 2.29 |
| 7H-benzo[c]carbazole | 194-59-2 | 4.72 | 138 | 408 | 0.3 | $9.0 \times 10^{-8}$ | $1.70 \times 10^{-6}$ | 9.91 | 4.99 | 5760 | $5.2 \times 10^{4}$ | 2.87 | 2.29 |
| Dibenz[c,h]acridine | 224-42-0 | 5.90 | 219 | 471 | 0.025 | $3.5 \times 10^{-9}$ | $8.50 \times 10^{-8}$ | 11.23 | 6.1 | 17520 | $1.6 \times 10^{5}$ | 3.72 | 2.09 |
| 13(H)-Dibenzo[a,i]carbazole | 197-00-2 | 6.10 | 254 | 480 | 0.01 | $1.8 \times 10^{-9}$ | $7.20 \times 10^{-8}$ | 11.6 | 6.3 | 17520 | $1.6 \times 10^{5}$ | 3.80 | 2.07 |
| 4-Methylbenzothiophene | 1455-18-1 | 3.62 | 14 | 259 | 26 | $8.5 \times 10^{-3}$ | $1.70 \times 10^{-5}$ | 6.33 | 3.25 | 900 | $8.1 \times 10^{3}$ | 2.12 | 2.08 |
| 2,3-Dimethylbenzothiophene | 1964-44-9 | 4.10 | 50 | 268 | 7.2 | $3.2 \times 10^{-3}$ | $2.60 \times 10^{-5}$ | 6.82 | 3.67 | 360 | $3.2 \times 10^{3}$ | 2.40 | 2.20 |
| 2,3,4-Trimethylbenzothiophene | 16587-45-4 | 4.59 | 46 | 285 | 2.3 | $1.0 \times 10^{-3}$ | $3.10 \times 10^{-5}$ | 7.34 | 4.13 | 360 | $3.2 \times 10^{3}$ | 2.68 | 2.32 |
| Benzothiophene | 95-15-8 | 3.12 | 32 | 221 | 130 | $1.8 \times 10^{-2}$ | $1.35 \times 10^{-5}$ | 5.83 | 2.93 | 360 | $3.2 \times 10^{3}$ | 1.80 | 1.94 |
| 2,3-Dibromothiophene | 3140-93-0 | 3.05 | -23 | 220 | 180 | $1.5 \times 10^{-2}$ | $4.20 \times 10^{-5}$ | 5.72 | 2.84 | 360 | $3.2 \times 10^{3}$ | 1.75 | 1.91 |
| 2-Bromobenzothiophene | 5118-13-8 | 3.83 | 41 | 261 | 12.5 | $5.5 \times 10^{-4}$ | $2.80 \times 10^{-5}$ | 7.21 | 3.6 | 1440 | $1.3 \times 10^{4}$ | 2.25 | 2.18 |
| 2,3-Dichlorobenzothiophene | 21948-62-9 | 4.06 | 58 | 278 | 6.2 | $2.8 \times 10^{-4}$ | $3.50 \times 10^{-5}$ | 7.57 | 3.82 | 1440 | $1.3 \times 10^{4}$ | 2.39 | 2.22 |
| 2,3-Dibromobenzothiophene | 5905-52-2 | 4.60 | 72 | 294 | 1.8 | $7.0 \times 10^{-5}$ | $4.20 \times 10^{-5}$ | 8.17 | 4.32 | 5760 | $5.2 \times 10^{4}$ | 2.68 | 2.32 |
| Dibenzothiophene | 132-65-0 | 4.38 | 97 | 333 | 1.47 | $2.1 \times 10^{-4}$ | $3.38 \times 10^{-5}$ | 7.7 | 4.27 | 1440 | $1.3 \times 10^{4}$ | 2.60 | 2.38 |
| 4-Methyldibenzothiophene | 7372-88-5 | 4.92 | 75 | 347 | 0.45 | $6.5 \times 10^{-5}$ | $4.10 \times 10^{-5}$ | 8.25 | 4.77 | 5760 | $5.2 \times 10^{4}$ | 2.91 | 2.28 |
| 2,4-Dimethyldibenzothiophene | 31317-59-2 | 5.45 | 80 | 360 | 0.14 | $2.0 \times 10^{-5}$ | $5.00 \times 10^{-5}$ | 8.79 | 5.27 | 5760 | $5.2 \times 10^{4}$ | 3.18 | 2.17 |
| 2,4,7-Trimethyldibenzothiophene | 31317-63-8 | 5.98 | 87 | 375 | 0.043 | $6.0 \times 10^{-5}$ | $6.00 \times 10^{-5}$ | 9.34 | 5.77 | 8760 | $7.9 \times 10^{4}$ | 3.45 | 2.11 |
| 2-Bromodibenzothiophene | 4895-05-0 | 5.10 | 105 | 352 | 0.24 | $3.5 \times 10^{-5}$ | $4.30 \times 10^{-5}$ | 8.5 | 4.94 | 5760 | $5.2 \times 10^{4}$ | 3.00 | 2.25 |
| 2,8-Dibromodibenzothiophene | 22796-85-6 | 5.82 | 128 | 392 | 0.038 | $5.0 \times 10^{-6}$ | $5.10 \times 10^{-5}$ | 9.34 | 5.63 | 8760 | $7.9 \times 10^{4}$ | 3.40 | 2.13 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Benzo[*b*]naphtho[1,2-*d*]thiophene | 239-35-0 | 5.80 | 121 | 430 | 0.056 | $5.0 \times 10^{-7}$ | $5.20 \times 10^{-6}$ | 9.95 | 5.75 | 8760 | $7.9 \times 10^{4}$ | 3.40 | 2.13 |
| Benzo[*b*]naphtho[2,3-*d*]thiophene | 243-46-9 | 5.80 | 148 | 432 | 0.052 | $4.8 \times 10^{-7}$ | $5.30 \times 10^{-6}$ | 9.97 | 5.75 | 8760 | $7.9 \times 10^{4}$ | 3.40 | 2.13 |
| 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | 66197-18-0 | 5.36 | 101 | 412 | 0.18 | $2.8 \times 10^{-6}$ | $1.40 \times 10^{-5}$ | 9.07 | 5.32 | 5760 | $5.2 \times 10^{4}$ | 3.17 | 2.18 |
| 6-Methylbenzo[b]naphtho[2,1-d]thiophene | 66197-22-6 | 6.34 | 112 | 445 | 0.018 | $1.5 \times 10^{-7}$ | $4.80 \times 10^{-6}$ | 10.52 | 6.25 | 17520 | $1.6 \times 10^{5}$ | 3.70 | 2.10 |
| Dibenzofuran | 132-64-9 | 4.12 | 83 | 287 | 3.1 | $2.5 \times 10^{-3}$ | $5.82 \times 10^{-5}$ | 6.97 | 3.75 | 1440 | $1.3 \times 10^{4}$ | 2.46 | 2.49 |
| 2-Methyldibenzofuran | 3846-00-2 | 4.66 | 54 | 300 | 0.88 | $7.2 \times 10^{-4}$ | $7.10 \times 10^{-5}$ | 7.52 | 4.26 | 5760 | $5.2 \times 10^{4}$ | 2.76 | 2.35 |
| Benzo[b]naphtho[2,3-d]furan | 243-17-4 | 5.65 | 152 | 390 | 0.078 | $6.0 \times 10^{-7}$ | $4.60 \times 10^{-6}$ | 9.63 | 5.57 | 8760 | $7.9 \times 10^{4}$ | 3.30 | 2.15 |
| Dinaphtho[1,2-b;1',2'-d]furan | 193-43-1 | 7.15 | 204 | 486 | 0.0015 | $2.0 \times 10^{-9}$ | $9.50 \times 10^{-8}$ | 11.93 | 6.8 | 17520 | $1.6 \times 10^{5}$ | 4.10 | 2.04 |

a – octanol-water partition coefficient, b – melting point, c – boiling point, d – octanol-air partition coefficient, e – soil adsorption coefficient, f – bioconcentration factor, and g – biomagnification factor (The values were obtained from US EPA ECOSAR database https://episuite.dev/#/)

## 1.4. ENVIRONMENTAL SOURCES OF HETEROCYCLIC AROMATIC COMPOUNDS

Source identification of environmental pollutants is essential for effective risk assessment and the development of targeted mitigation and regulatory strategies. As mentioned earlier, all classes of PACs, including HPACs, generally originate from three primary sources: pyrogenic, petrogenic, and biogenic processes.[57] These origins can be either natural or anthropogenic. In addition to these primary origins, HPACs can also be formed through secondary transformation processes such as photochemical reactions, microbial activity, or interactions with atmospheric oxidants.[58,59]

Pyrogenic processes are among the dominant sources of HPACs across most environmental matrices. These compounds form during the incomplete combustion of organic matter in the presence of nitrogen, sulfur, or oxygen-containing species. Pyrogenic processes primarily produce unsubstituted parent HPACs such as carbazole, quinoline, dibenzothiophene, and benzofuran.[60–62]

Anthropogenic pyrogenic processes that form HPACs include vehicular emissions, wood burning, and coal combustion, as well as natural pyrogenic processes such as wildfires and volcanic activity. Studies confirm the formation of carbazole, oxygenated PAHs, and dibenzothiophene in gasoline engine exhaust.[63] Coal combustion is also found to emit alkylated carbazoles and benzocarbazoles.[64] Another interesting pyrogenic PAC source is food processing. Smoking and barbecuing food using wood and charcoal have been found to form HPACs.[65–67]

Petrogenic processes are another major source of PACs; these PACs form during rock formation and petroleum production. Therefore, unburnt petroleum and crude oil can release large amounts of alkylated HPACs to the environment via natural seepage, spills, leaks, and during refining activities.[68–70] These compounds have been identified in coal-based products (coal tar) and petroleum-based products (diesel and asphalt).[70,71] Compound mixtures originating from these two sources have been shown to exhibit distinct HPAC profiles. For instance, carbazole levels in coal tar are much higher than in asphalt-based products.[72] These differences are very important in source identification studies.

Heterocyclic compounds are a significant class of compounds found in biological systems. These compounds play essential structural and functional roles in those systems. Macromolecules such as DNA, RNA, and proteins, as well as amino acids such as tryptophan, are examples. These naturally occurring compounds act as biogenic precursors to HPACs and other types of PACs. During natural processes such as biomass burning, diagenesis, and thermal alteration of organic matter, these biologically derived compounds can transform into HPACs.[73] For example, indole can be formed from microbial transformation of tryptophan and end up in biological waste.[74] Although biogenic sources are generally less dominant than pyrogenic and petrogenic inputs, they are also important in source identification studies.

In addition to these three major processes, HPACs are also produced by industrial activities such as textile and pharmaceutical production.[75,76] Derivatives of quinoline, indole, and acridine type PANHs are used in the pharmaceutical industry, and this industrial waste may contain these compounds.[77,78] Also, quinolines and halogenated indigo dyes are found in the textile industry.[76,79]

Source characterization studies are an essential part of environmental pollutant analysis. Any source identification study depends on three main prerequisites: (i) the presence of unique fingerprints across different sources, (ii) the chemical persistence of fingerprints in the sink, and, (iii) the use of appropriate analytical or statistical methods.[80] Many studies have used PAHs as source fingerprints for PACs, however, in contaminated environmental samples, the PAH ratios remain in similar ranges, and unique fingerprints are not available.[81–83] The use of HPACs as chemical fingerprints is much more reliable because of their low concentrations and distinct profiles across different sources.[72] However, persistence studies of HPACs are limited, and that hinders the possibility of using HPACs as chemical fingerprints in different environmental matrices.[84]

## 1.5 ENVIRONMENTAL DISTRIBUTION OF HETERO-POLYCYCLIC AROMATIC COMPOUNDS

Hetero-polycyclic aromatic compounds are widely distributed in environmental matrices such as air, water, soil, sediments, and biota. Even though these compounds share origins with the parent

PAHs, the presence of heteroatoms makes their environmental behavior different from that of their homocyclic counterparts. As previously discussed, due to differences in p-chem properties such as polarity, reactivity, and partition coefficients, HPACs may exhibit distinct distribution patterns in the environment.

The presence of HPACs in environmental matrices was first reported in the early 1980s, most likely due to advances in gas chromatography mass spectrometry (GC-MS) techniques.[85,86] Although HPACs were detected in early studies, they were typically present as minor constituents. This may have contributed to the limited research attention these compounds received, mainly due to the analytical challenges associated with their detection and characterization. Concurrent with this, regulatory and scientific focus became heavily centered on PAHs following the prioritization of the 16 parent PAHs as priority pollutants by the USA-EPA.[4] This classification led to extensive research into the environmental fate, transport, and toxicity of PAHs, establishing them as key indicators of contamination.[87–89] Consequently, HPACs were often overlooked despite emerging evidence of their distinct environmental behavior and toxicological profiles.[12]

Just as other subclasses of PACs, HPACs can also be found in abiotic (e.g., air, soil, sediment, dust, and snow) and biotic matrices. Table 1.3 summarizes some studies that detected at least one type of HPACs in these matrices. Several studies have determined HPAC concentrations in air and airborne particles. These studies highlight the origin of these compounds from combustion sources

associated with asphalt mixing, diesel exhaust, and oil extraction activities.[89–90] These studies show that detectable HPAC concentrations are often comparable to or higher than those of PAHs, underscoring the importance of regular monitoring of HPACs in air samples.[90,91] Also, Manzano *et al*.[92] and Chibwe et al.[64] confirm that petroleum coke is one of the main sources of PANHs and HMW compounds, such as alkylated carbazoles and benzocarbazoles, which have the potential to undergo long-range transport from the source.

Due to the nonpolar nature of PACs, sediments are more likely to be the primary repository for these compounds introduced into aquatic systems. Compared to other types of matrices, many studies have been done to identify HPACs in sediment samples.[18] However, most of these studies analyze one or two classes or a small number of HPACs as part of PAH studies. Dibenzothiophenes are among the most characterized HPACs in sediments, alongside homocyclic PACs.[93,94] These studies demonstrate persistent contamination of sediments by HPACs from oil sands operations, urban and industrial runoff, and coal tar, with bioaccumulation, long-range transport, and elevated toxicity risks.[64,95,96]

There are also several studies that focus on measuring HPACs in soil samples.[97–99] Besides the previously mentioned analytical limitations associated with HPAC analysis, an additional major challenge arises when dealing with complex environmental matrices such as sediments and soils. In these systems, analytes are often strongly bound to organic matter, black carbon, and mineral

surfaces, making exhaustive extraction difficult.[18] Several of these studies highlight that there are also many challenges when it comes to the remediation of these contaminants from these matrices.[100–103]

In addition, HPACs have been identified in various biological species, including fish and benthic organisms, indicating their potential for bioaccumulation and persistence.[104] However, the number of studies that focus on HPACs is much lower than that of other classes of PACs. Available evidence suggests that biota samples are exposed to HPACs through multiple pathways, including air, water, and dietary uptake. These exposure routes can enhance accumulation, particularly at higher trophic levels, indicating potential trophic transfer and food web magnification and raising concerns about ecological risk.[105,106]

Compared with non-halogenated HPACs, research on halogenated HPACs in environmental samples is very limited. Studies to date have focused on polyhalogenated compounds rather than HPACs with fewer halogen substitutes. For example, polyhalogenated dibenzofurans and polyhalogenated carbazoles have been previously studied in sediments[107–109] and biota samples.[48,109,110]

The analysis of HPACs in environmental samples has relied on a range of extraction and detection techniques tailored to complex matrices. Conventional extraction methods, including Soxhlet extraction, as well as newer techniques such as accelerated solvent extraction (ASE) and

QuEChERS, are commonly employed. [111] However, these approaches are often solvent-intensive and time-consuming. They may suffer from limited selectivity, leading to the co-extraction of interfering matrix components such as lipids, organic matter, and sulfur-containing compounds. Because of the wide polarity range of HPACs, achieving efficient and consistent extraction recoveries across compound classes is challenging. The available techniques have been used with nonpolar solvent systems, plus the addition of polar solvents like DCM: acetone mixtures to extract HPACs.[91]

Detection and identification of HPACs have primarily been carried out using GC-MS. Also, gas chromatography coupled with selective detectors like an atomic emission detector and a diode array detector [90] were also used in some studies. Gas chromatography coupled with tandem mass spectrometry (GC-MS/MS) is one of the most reliable techniques used for HPACs analysis. This is mainly because of the high selectivity of multiple reaction monitoring (MRM) methods of GC-MS/MS, compared to less selected ion monitoring (SIM). [112]

While GC-based methods provide high sensitivity and separation efficiency for nonpolar and semi-volatile compounds, they are less suitable for more polar or thermally labile HPACs without derivatization. Furthermore, co-elution with structurally similar PAHs and APAHs remains a significant challenge, particularly in one-dimensional GC systems and low-resolution mass spectrometric techniques.[113,114]

Hence, more advanced techniques such as comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry (GC×GC–TOFMS) and other high-resolution mass spectrometry techniques, such as Orbitrap mass spectrometry, are needed.[64,113] Although high-resolution and multidimensional techniques have improved compound separation and identification, they often require complex data processing workflows and still depend heavily on spectral libraries, which are limited for many HPACs.[115] As a result, accurate identification and quantification remain constrained for HPACs. In my thesis, I have used a GC-MS/MS for targeted analysis of PACs and a GC×GC–TOFMS for non targeted analysis of PACs.

**Table 1.3:** The environmental distribution data of hetero-polycyclic aromatic compounds

| Compounds Analysed | Location | Matrix Type | Extraction Techniques | Detection Techniques | Concentration Range | Year | Reference Name |
|---|---|---|---|---|---|---|---|
| PASHs (53 total) | Aluminum melting facility | Airborne particulates | Soxhlet extraction (DCM, 15h) | GC-AED, GC-MS | N/A | N/A | Becker et al.[116] |
| PANHs, PASHs, PAOHs (54 total, 31 detected) | Brazilian refineries | Asphalt fumes | A fume generator with an Argon flow | HPLC-DAD | N/A (Temp-dependent) | 2024 | Cargnin et al.[90] |
| PASHs (as a part of APAH study) | Athabasca Oil Sands Region (AOSR), Alberta, Canada | Ambient air | Soxhlet extraction (DCM: acetone – 10:1) | GC-MS (SIM) | Median ~12 ng/m³ total DBTs | 2011–2015 | Wnorowski et al.[91] |
| PASHs | Cement industry areas, Central Italy | Honeybees (as a proxy of air) | Sonication (DCM, 10 min) | GC-MS/MS with an iron trap detector | 1.55–35.63 ng/g (dw) | 2023 | Sebastiani et al.[117] |
| PANHs (>200) | AOSR, Canada | air, snow, lake sediment, petcoke | Solvent extraction (air, petcoke)<br>ASE (snow filters)<br>Soxhlet extraction (sediments) | GC×GC-TOFMS and LC-Orbitrap-MS | 0.311–10100 pg/L at reference site for snow; decrease with distance/depth for all | 2011–2014 | Manzano et al.[92] |
| PANHs (>200) | AOSR, Canada | Air, snow, sediment | Solvent extraction (air)<br>ASE (snow filters)<br>Soxhlet extraction (sediments) | GC×GC-TOFMS and LC-Orbitrap-MS | 0.311–10100 pg/L (snow); decrease with distance/depth | 2014–2015 | Chibwe et al.[64] |
| Indole and carbazole (PANHs) | Great Lakes, USA | Surface water and sediment | ASE | HPLC-MS/MS | 0.0284 µg/L (water); 130 µg/kg (sediment) | 2013–2014 | Elliott et al.[118] |
| PANHs, PASHs, and PAOHs (4 total) | Pennsylvania, USA | Urban stream sediments | ASE | GC-MS | 59–1897 µg/kg | 2010 | Witter & Nguyen[119] |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| PANHs | Austin, Texas, USA | Pavement runoff/sediment | Liquid- liquid extraction<br>Solid phase extraction | GC-MS | ~1 – 750 µg/L | 2011 | Mahler et al.[70] |
| PANHs | Southeast Australia | Sediment, Rock oyster | Vortex + ultrasonication (Sediment)<br>QuEChERS (Rock Oyster) | GC×GC-TOFMS | Mean 7.18 µg/ g (Sediment)<br>~3 - 5 µg/g (Rock oyster) | Not specified | Idowu et al.[111] |
| PASHs | Sichuan Basin, China | Soil and sediment | Soxhlet extraction (DCM) | GC-MS | median 2327.2 ng/g | N/A | Song et al.[120] |
| PANHs, PASHs, and PAOHs (15 total) | Wyckoff/ Eagle Harbor, USA | Soil | ASE (DCM) | GC-MS | ΣPANHs: 18 µg/g<br>ΣPASHs: 7.2 µg/g<br>ΣPAOHs: 4.5 µg/g | 2022 | Titaley et al.[121] |
| PANHs and PASHs | Newcastle, Australia | Soils (industrial heritage city) | ASE (Hexane: acetone – 3:1) | GC-MS (SIM) | 15 – 6240 ng/ g | N/A | Idowu et al.[122] |
| Oxy-PAHs (10), N-PACs (4) | Contaminated soils | Soil | Pressurized liquid extraction | GC-MS | Varied | 2014 | Lundstedt et al.[123] |
| Quinolines (PANHs) | Stockholm, Sweden | Textiles (proxy dust/soil) | Ultrasound-assisted solvent extraction | GC-MS | Up to 1.9 mg/garment | 2011–2012 | Luongo et al.[76] |
| PASHs | AOSR, Canada | Lichen | Ultrasonic bath | GC-MS | 6- 69 ng/g | 2019 | Landis et al.[81] |
| PANHs, PASHs, and PAOHs | German tar sites | Groundwater | Liquid-liquid extraction | GC-MS | <LOD–8071 µg/L | 2008 | Schlanges et al.[124] |
| PASHs | AOSR, Canada | Granulosa cells (in vitro) | - | GC-MS/MS | 0.1–10 nM | 2025 | Perono et al.[112] |
| PASHs | Arctic Canada (Baffin Bay-Davis Strait) | Seabird tissues | ASE (DCM) | GC-MS/MS | 3.7–40.7 ng/g lw | 2007-2008, 2018 | Provencher et al.[105] |
| Halogenated Carbazoles | The Great Lakes | Sediments | ASE | GC-MS/MS | 4 – 40 ng/g | 2013 | Li et al.[107] |

## 1.6 NEGATIVE IMPACTS OF HETERO-POLYCYCLIC AROMATIC COMPOUNDS ON HUMANS AND OTHER BIOGENIC ENVIRONMENT

### 1.6.1 Toxic properties of hetero-polycyclic aromatic compounds

The toxic effects of PACs have been studied for more than a century. The identification of benzo[a]pyrene as a potent carcinogen in experimental studies during the 1930s marked a critical milestone in research on chemical carcinogenesis.[125] Subsequent evaluations by the International Agency for Research on Cancer (IARC) in the 1970s classified benzo[a]pyrene as a carcinogenic compound, and this further solidified its role as a reference compound in toxicological studies.[126] Since then, scientific interest in PACs has grown substantially, driven by their widespread environmental occurrence, persistence, and diverse toxicological effects.

Polycyclic aromatic compounds exhibit a range of toxicities, often mediated by the aryl hydrocarbon receptor (AhR) pathway, leading to CYP1A induction and oxidative stress.[127] Similar to PAHs, the toxicity of HPACs has endpoints such as acute toxicity, genotoxicity, carcinogenicity, and teratogenicity.[12] Table 1.4 summarizes studies that have identified the toxicity effects of HPACs. Overall, the toxic properties of HPACs depend on their structures, mainly on the type and position of heteroatoms.

The main toxic mechanisms of HPACs include metabolic activation leading to reactive intermediates, DNA adduct formation, reactive oxygen species (ROS) formation, phototoxicity, and protein adduct formation leading to enzyme dysfunction. Bioactivation is the central mechanism of toxicity for most PACs. In biological systems, these compounds are transformed into more toxic species (bioactivated) during Phase I metabolic pathways. This is primarily mediated by xenobiotic-metabolizing enzymes, such as cytochrome P450S.[128]

Major metabolic pathways of bioactivation are the dihydrodiol epoxide pathway, the radical cation pathway, and the ortho-quinone pathway.[129] The metabolites produced during these pathways can covalently bind to DNA, forming DNA adducts, bind to proteins and enzymes, and generate ROS. These processes cause toxicity effects in humans and other species.[130] The DNA adducts formed can cause mutations leading to cancer. Also, some PANHs form hydroxyl derivatives that form nitrenium ions, which can bind to guanine bases in DNA, forming stable adducts.[131] Also, oxidative DNA damage can be caused by ROS.[129] All these DNA changes can cause HPAC-mediated genotoxicity.

Heterocyclic compounds (e.g., quinones) undergo redox cycling, generating ROS such as superoxide radicals and hydrogen peroxide. These reactive species can damage lipids, oxidize proteins, and cause oxidative DNA lesions. Oxidative stress is therefore a major contributor to both acute toxicity and long-term effects such as cancer and neurotoxicity.[129]

These compounds can also act as endocrine disruptors, interfering with hormone signaling pathways. This is a receptor-mediated endocrine disruption primarily involving the AhR and nuclear hormone receptors, such as the estrogen receptor (ER) and thyroid hormone receptors. This causes altered gene expression, hormonal imbalances, and adverse effects on development, reproduction, and metabolic regulation.[128,132–134]

Certain HPACs exhibit phototoxicity, in which exposure to UV light enhances their toxicity. Upon irradiation, these compounds form excited-state intermediates that generate ROS, increasing cellular damage.[12] These toxic effects confirm that regular detection of PACs is essential to ensure environmental and human health.

**Table 1.4:** Experimental evidence on toxic effects of hetero-polycyclic aromatic compounds

| **Toxic Effect** | **Example HPACs** | **Reference** |
|---|---|---|
| Potential adverse outcomes (e.g., tumors, skewed sex ratios, reproductive dysfunction) | Indole, carbazole | Baldwin et al.[135] |
| Mutagenicity, carcinogenicity | PANHs | Barzegara et al.[136] |
| Photoenhanced toxicity | PANHs | Bleeker et al.[137] |
| Chronic reproductive toxicity, growth inhibition, locomotor changes (hyperactivity, spinning) | 7H-benzo[c]carbazole), benzo[b]naphtho[1,2-d]furan, benzo[b]naphtho[1,2-d]thiophene, dinaphtho[2,1-b:1′,2′-d]furan | Çelik et al.[49] |
| Developmental toxicity, AHR activation, oxidative stress | Carbazole, acridine, dibenzofuran, 5,6-benzoquinoline, indole, quinoline, thianaphthene | Chlebowski et al.[138] |
| Developmental toxicity, mutagenicity | Benzo[a]carbazole | Dutson et al.[137] |
| Acute/chronic toxicity, genotoxicity, carcinogenicity, endocrine disruption, phototoxicity | PANHs, PASHs and PAOHs | Ghosh et al.[12] |
| Greater genotoxicity, mutagenicity, carcinogenicity, and developmental toxicity than parent PAHs | PANHs, PASHs and PAOHs | Idowu et al.[6] |
| Chronic fish toxicity | PASHs | Koleva et al.[20] |
| Genotoxicity (DNA damage), developmental toxicity (abnormalities, reduced growth) and embryotoxicity | PASHs | Le Bihanic et al.[139] |
| Chronic toxicity to invertebrates (reduced survival/reproduction) | PANHs | Paumen et al.[140] |
| Embryotoxicity | PANHs and PAOHs | Peddinghaus et al.[141] |
| Cytotoxicity and AhR activation | Dibenz[a,h]acridine, dibenz[a,i]acridine, dibenz[a,j]acridine, benz[a]acridine, benz[c]acridine | Sovadinová et al.[142] |
| Embryotoxicity, CYP1A inhibition, heart elongation, and pericardial edema | Carbazole and dibenzothiophene | Wassenberg et al.[143] |

## 1.7 SCOPE OF MY RESEARCH

My primary research aim is to elucidate the fate, distribution and behaviour of HPACs in the environment. To achieve this, I developed and validated analytical methods for the quantitative analysis of HPACs in environmental samples using GC-MS/MS. I subsequently applied advanced chemometric techniques to characterize their environmental fate and behavior. In addition, comprehensive two-dimensional GC coupled with high-resolution time-of-flight mass spectrometry (GC x GC-HRTOF-MS) was employed to identify non-targeted PACs in the environmental samples.

## 1.8 RESEARCH HYPOTHESES

The hypotheses assessed in my thesis are as follows:

*Chapter 03:* HPACs are present in the Lake Ontario sediments.

*Chapter 04:* HPACs are more efficient source of fingerprints of PACs than the 16 PAHs.

*Chapter 05:* HPACs are present in Lake Erie fish, and their concentration profiles can vary by basin/location.

*Chapter 06:* Bioaccumulation patterns of HPACs are species-dependent and are affected by environmental parameters.

*Chapter 07:* Highlighting the importance of integration of HPACs in environmental monitoring results in a more accurate interpretation of PACs.

*Chapter 08:* Development of data processing filters for PACs detected in non-targeted analysis improves the accuracy of non-targeted data.

## 1.9 THESIS STRUCTURE

My thesis is divided into nine (9) chapters.

**Chapter 1** provides an overview of the properties, occurrence, and toxicity of the analytes of interest (HPACs), as well as my thesis hypotheses and scope. **Chapter 2** includes brief descriptions of the methodology for sample processing and instrumental analysis, statistical and chemometric analyses, and non-targeted data processing methods used in this study. **Chapter 3** discusses GC-MS/MS methods and the development and validation of extraction and purification methods for HPACs in sediments. **Chapter 4** demonstrates how HPACs can serve as source fingerprints for PACs, and the identified HPACs were used to fingerprint Lake Ontario sediments. **Chapter 5** presents how the HPACs and other PAC concentrations vary in fish samples collected from different basins of Lake Erie. **Chapter 6** focuses on the bioaccumulation of HPACs in the Lake Erie food web. **Chapter 7** shows that HPACs can serve as the most accurate indicators of the environmental behavior of PACs. **Chapter 8** shows how new data-filtering methods can ensure accurate identification of non-targeted PAC data. **Chapter 9** gives an overall conclusion of the study. Suggestions and recommendations for future research in relation to the application of this work are also mentioned.

## 1.10 REFERENCES


(1) Crude Oil Industry Overview. https://natural-resources.canada.ca/energy-sources/fossil-fuels/crude-oil-industry-overview?utm_source=chatgpt.com. Accessed on May 15, 2026.

(2) Wang, W. The Oil and Gas Sector in Canada: A Year after the Start of the Pandemic Economic and Social Reports. https://doi.org/10.25318/36280001202100700003-eng.

(3) Lee, Kenneth.; Boufadel, Michel.; Chen, Bing. The Royal Society of Canada Expert Panel : The Behaviour and Environmental Impacts of Crude Oil Released into Aqueous Environments; Royal Society of Canada, **2015**.

(4) Achten, C.; Andersson, J. T. Overview of Polycyclic Aromatic Compounds (PAC). *Polycycl. Aromat. Compd*. **2015**, 35 (2–4), 177–186. https://doi.org/10.1080/10406638.2014.994071.

(5) Suresh, A.; Soman, V.; K R, A.; A, R.; Rahman K, H. Sources, Toxicity, Fate and Transport of Polyaromatic Hydrocarbons (PAHs) in the Aquatic Environment: A Review. *Environmental Forensics*. Taylor and Francis Ltd. **2025**, pp 35–57. https://doi.org/10.1080/15275922.2024.2366801.

(6) Idowu, O.; Semple, K. T.; Ramadass, K.; O'Connor, W.; Hansbro, P.; Thavamani, P. Beyond the Obvious: Environmental Health Implications of Polar Polycyclic Aromatic Hydrocarbons. *Environment International*. Elsevier Ltd February 1, **2019**, pp 543–557. https://doi.org/10.1016/j.envint.2018.12.051.

(7) Xia, Z.; Idowu, I.; Marvin, C.; Thomas, P. J.; Johnson, W.; Francisco, O.; Stetefeld, J.; Crimmins, B.; Fry, M.; Tomy, G. T. Identification of Halogenated Polycyclic Aromatic Hydrocarbons in Biological Samples from Alberta Oil-Sands Region. *Chemosphere* **2019**, 215, 206–213. https://doi.org/10.1016/j.chemosphere.2018.10.050.

(8) Wallace, S. J.; de Solla, S. R.; Head, J. A.; Hodson, P. V.; Parrott, J. L.; Thomas, P. J.; Berthiaume, A.; Langlois, V. S. Polycyclic Aromatic Compounds (PACs) in the Canadian Environment: Exposure and Effects on Wildlife. *Environmental Pollution*. Elsevier Ltd October 1, **2020**. https://doi.org/10.1016/j.envpol.2020.114863.

(9) Toxic substances list: schedule 1. Canadian Environmental Protection Act 1999, https://www.canada.ca/en/environment-climate-change/services/canadian-environmental-protection-act-registry/substances-list/toxic/schedule-1.html.

(10) Andersson, J. T.; Achten, C. Time to Say Goodbye to the 16 EPA PAHs? Toward an Up-to-Date Use of PACs for Environmental Purposes. *Polycycl. Aromat. Compd*. **2015**, 35 (2–4), 330–354. https://doi.org/10.1080/10406638.2014.991042.

(11) Marvin, C. H.; Berthiaume, A.; Burniston, D. A.; Chibwe, L.; Dove, A.; Evans, M.; Hewitt, L. M.; Hodson, P. V.; Muir, D. C. G.; Parrott, J.; Thomas, P. J.; Tomy, G. T. Polycyclic Aromatic

Compounds in the Canadian Environment: Aquatic and Terrestrial Environments. *Environmental Pollution.* Elsevier Ltd September 15, **2021**. https://doi.org/10.1016/j.envpol.2021.117442.

(12) Ghosh, P.; Mukherji, S. Fate, Detection Technologies and Toxicity of Heterocyclic PAHs in the Aquatic and Soil Environments. *Science of the Total Environment*. Elsevier B.V. September 20, **2023**. https://doi.org/10.1016/j.scitotenv.2023.164499.

(13) Marvin, C. H.; Tomy, G. T.; Thomas, P. J.; Holloway, A. C.; Sandau, C. D.; Idowu, I.; Xia, Z. Considerations for Prioritization of Polycyclic Aromatic Compounds as Environmental Contaminants. *Environmental Science and Technology.* American Chemical Society December 1, **2020**, pp 14787–14789. https://doi.org/10.1021/acs.est.0c04892.

(14) Balaban, A. T.; Oniciu, D. C.; Katritzky, A. R. Aromaticity as a Cornerstone of Heterocyclic Chemistry. *Chem. Rev*. **2004**, 104 (5), 2777–2812. https://doi.org/10.1021/cr0306790.

(15) Posthuma, J. The composition of petroleum; **1977**; Vol. 171.

(16) SNYDER VOl, L. R.; Arnold, R. L.; Biranowski, J. B.; Carlson, iW W.; R Corjield, P. W.; Doomes, E.; Frame, R. R.; Hemwall, R. W.; Hoyt, E. B.; Jarvis, B. B.; Mecca, T. G.; Schexnayder, D. A.; Scamehorn, R. G.; Springer, W. A.; Strong, J. G.; Vestling, M. M.; Williams, J. M.; Yee, K. C.; Snyder, L. R. Petroleum Nitrogen Compounds and Oxygen Compounds; **1970**.

(17) Comber, M. H. I.; De Ferrer, J. A.; Djemel, N.; Eadsforth, C. V; Lescrauwaet, A.; Léon, M.; Linington, P. S.; Redman, A.; Villalobos, S. A.; Den Haan, K. H. Analysis of N-, O-, and S-Heterocyclics in Petroleum Products Using GCxGC with Specific Detection; **2016**. www.concawe.org.

(18) Machado, M. E.; Nascimento, M. M.; Bomfim Bahia, P. V.; Martinez, S. T.; Bittencourt de Andrade, J. Analytical Advances and Challenges for the Determination of Heterocyclic Aromatic Compounds (NSO-HET) in Sediment: A Review. *TrAC - Trends in Analytical Chemistry*. Elsevier B.V. May 1, **2022**. https://doi.org/10.1016/j.trac.2022.116586.

(19) Pais, P.; Knize, M. G. Chromatographic and Related Techniques for the Determination of Aromatic Heterocyclic Amines in Foods; **2000**; Vol. 747. www.elsevier.com/locate/chromb.

(20) Koleva, Y.; Tasheva, Y. The persistence, bioaccumulation and toxic estimation of some sulfur compounds in the environment. *Petroleum and Coal*, **2012** 54, pp 220 – 224.

(21) Salam, L. B.; Ilori, M. O.; Amund, O. O. Properties, Environmental Fate and Biodegradation of Carbazole. 3 *Biotech.* Springer Verlag June 1, **2017**. https://doi.org/10.1007/s13205-017-0743-4.

(22) Bioactive Heterocycles V; Mahmud Tareq Hassan Khan, Ed.; Springer Berlin, Heidelberg, **2007**.

(23) Alvárez-Builla, J.; Barluenga, J. Heterocyclic Compounds: An Introduction. In Modern Heterocyclic Chemistry; Wiley-VCH, **2011**; Vol. 1, pp 1–9. https://doi.org/10.1002/9783527637737.ch1.

(24) Fan, X. L.; Wang, X. Q.; Wang, J. T.; Li, H. D. Structural and Electronic Properties of Linear and Angular Polycyclic Aromatic Hydrocarbons. Physics Letters, Section A: General, Atomic and Solid State Physics **2014**, 378 (20), 1379–1382. https://doi.org/10.1016/j.physleta.2014.03.035.

(25) Matta, C. F.; Hernández-Trujillo, J. Bonding in Polycyclic Aromatic Hydrocarbons in Terms of the Electron Density and of Electron Delocalization. *Journal of Physical Chemistry A* **2003**, 107 (38), 7496–7504. https://doi.org/10.1021/jp034952d.

(26) Poater, J.; Fradera, X.; Duran, M.; Solà, M. The Delocalization Index as an Electronic Aromaticity Criterion: Application to a Series of Planar Polycyclic Aromatic Hydrocarbons. *Chemistry - A European Journal* **2003**, 9 (2), 400–406. https://doi.org/10.1002/chem.200390041.

(27) Cordell, F. R.; Boggs, J. E. Structure and degree of aromatic character in furan, pyrrole, and thiophene. J. *Mol. Struct*. **1981**, 85, 163–178.

(28) Ghomri, A.; Mekelleche, S. M. Reactivity and Regioselectivity of Five-Membered Heterocycles in Electrophilic Aromatic Substitution: A Theoretical Investigation. Journal of Molecular Structure: THEOCHEM **2010**, 941 (1–3), 36–40. https://doi.org/10.1016/j.theochem.2009.10.035.

(29) Galabov, B.; Ilieva, S.; Cheshmedzhieva, D.; Nikolova, V.; Popov, V. A.; Hadjieva, B.; Schaefer, H. F. Mini-Review on Structure-Reactivity Relationship for Aromatic Molecules: Recent Advances. *ACS Omega*. American Chemical Society March 15, **2022**, pp 8199–8208. https://doi.org/10.1021/acsomega.1c07176.

(30) Kromann, J. C.; Jensen, J. H.; Kruszyk, M.; Jessing, M.; Jørgensen, M. Fast and Accurate Prediction of the Regioselectivity of Electrophilic Aromatic Substitution Reactions. *Chem. Sci*. **2018**, 9 (3), 660–665. https://doi.org/10.1039/c7sc04156j.

(31) Szatylowicz, H.; Stasyuk, O. A.; Krygowski, T. M. Calculating the Aromaticity of Heterocycles. *Adv. Heterocycl. Chem.* **2016**, 120, 301–327. https://doi.org/10.1016/bs.aihch.2016.03.007.

(32) Aromaticity in Heterocyclic Compounds; Tadeusz, M. K., Michał, K. C., Eds.; Springer Berlin, Heidelberg, **2009**.

(33) Ghosh, P.; Mukherji, S. Environmental Contamination by Heterocyclic Polynuclear Aromatic Hydrocarbons and Their Microbial Degradation. *Bioresource Technology*. Elsevier Ltd December 1, **2021**. https://doi.org/10.1016/j.biortech.2021.125860.

(34) Lim, J. M.; Shim, S.; Bui, H. T.; Kim, J.; Kim, H. J.; Hwa, Y.; Cho, S. Substitution Effect of a Single Nitrogen Atom on π-Electronic Systems of Linear Polycyclic Aromatic Hydrocarbons

(PAHs): Theoretically Visualized Coexistence of Mono- and Polycyclic π-Electron Delocalization. *Molecules* **2024**, 29 (4). https://doi.org/10.3390/molecules29040784.

(35) Li, X.; Staykov, A.; Yoshizawa, K. Orbital Views of the Electron Transport through Heterocyclic Aromatic Hydrocarbons. *Theor. Chem. Acc.* **2011**, 130 (4–6), 765–774. https://doi.org/10.1007/s00214-011-0968-y.

(36) Valencia, D.; Klimova, T.; García-Cruz, I. Aromaticity of Five- and Six-Membered Heterocycles Present in Crude Oils - An Electronic Description for Hydrotreatment Process. In Fuel; **2012**; Vol. 100, pp 177–185. https://doi.org/10.1016/j.fuel.2012.05.011.

(37) Mackay, D.; Shiu, Y. W.; Ma, K.-C.; Lee, S. C. Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals Volume I Introduction and Hydrocarbons, Second edition; **2006**.

(38) Maclagan, R. G. A. R.; Gronert, S.; Meot-Ner, M. Protonated Polycyclic Aromatic Nitrogen Heterocyclics: Proton Affinities, Polarizabilities, and Atomic and Ring Charges of 1-5-Ring Ions. *Journal of Physical Chemistry A* **2015**, 119 (1), 127–139. https://doi.org/10.1021/jp5069939.

(39) Nielsen, T.; Siigur, K.; Helweg, C.; Jørgensen, O.; Hansen, P. E.; Kirso, U. Sorption of Polycyclic Aromatic Compounds to Humic Acid as Studied by High-Performance Liquid Chromatography. *Environ. Sci. Technol*. **1997**, 31, 1102–1108.

(40) Waluk, J. Hydrogen-Bonding-Induced Phenomena in Bifunctional Heteroazaaromatics. *Acc. Chem. Res*. **2003**, 36 (11), 832–838. https://doi.org/10.1021/ar0200549.

(41) Tang, S.; Zhou, Z.; Jiang, Z.; Zhu, W.; Qiao, D. Indole-Based Tubulin Inhibitors: Binding Modes and SARs Investigations. *Molecules.* MDPI March 1, **2022**. https://doi.org/10.3390/molecules27051587.

(42) Saranjampour, P.; Vebrosky, E. N.; Armbrust, K. L. Salinity Impacts on Water Solubility and N-Octanol/Water Partition Coefficients of Selected Pesticides and Oil Constituents. *Environ. Toxicol. Chem.* **2017**, 36 (9), 2274–2280. https://doi.org/10.1002/etc.3784.

(43) Planeta, J.; Šťaváková, L.; Karásek, P.; Roth, M. Limiting Partition Coefficients of Sulfur-Containing Aromatics in a Biphasic [Bmim][MeSO4]-Supercritical CO2 System. *J. Chem. Eng. Data* **2011**, 56 (3), 527–531. https://doi.org/10.1021/je1010642.

(44) Friedrichsen, W. Recent Advances in the Chemistry of Benzo[c]Furans and Related Compounds; **1999**; Vol. 13.

(45) Shemer, H.; Linden, K. G. Aqueous Photodegradation and Toxicity of the Polycyclic Aromatic Hydrocarbons Fluorene, Dibenzofuran, and Dibenzothiophene. *Water Res.* **2007**, 41 (4), 853–861. https://doi.org/10.1016/j.watres.2006.11.022.

(46) Freitas, V. L. S.; Gomes, J. R. B.; Ribeiro Da Silva, M. D. M. C. Dibenzofuran and Methyldibenzofuran Derivatives: Assessment of Thermochemical Data. *Struct. Chem.* **2013**, 24 (6), 1923–1933. https://doi.org/10.1007/s11224-013-0272-7.

(47) Mackay, D.; Arnot, J. A.; Gobas, F. A. P. C.; Powell, D. E. Mathematical Relationships between Metrics of Chemical Bioaccumulation in Fish. *Environ. Toxicol. Chem.* **2013**, 32 (7), 1459–1466. https://doi.org/10.1002/etc.2205.

(48) Wu, Y.; Tan, H.; Zhou, C.; Crimmins, B. S.; Holsen, T. M.; Chen, D. Bioaccumulation and Spatiotemporal Trends of Polyhalogenated Carbazoles in Great Lakes Fish from 2004 to 2016. *Environ. Sci. Technol.* **2018**, 52 (8), 4536–4545. https://doi.org/10.1021/acs.est.8b00427.

(49) Çelik, G.; Healy, S. A.; Stolte, S.; Mayer, P.; Markiewicz, M. Daphnia Magna as an Alternative Model for (Simultaneous) Bioaccumulation and Chronic Toxicity Assessment─Controlled Exposure Study Indicates High Hazard of Heterocyclic PAHs. *Environ. Sci. Technol.* **2025**, 59 (18), 8984–8996. https://doi.org/10.1021/acs.est.5c00384.

(50) Picel ~, K. C.; Stamoudis, V. C.; Simmons, M. S. Distribution coefficients for chemical components of a coal-oil/water system. *Wat. Res* **1988**, 22 (9), 1189–1199.

(51) Dearden, J. C. Partitioning and Lipophilicity in Quantitative Structure-Activity Relationships; **1985**; Vol. 61. https://about.jstor.org/terms.

(52) Colmo, C.; Tegelaar, M.; Ayres, P. Bioremediation of Polycyclic Aromatic Hydrocarbons (PAHs) in Aqueous Environments: A Review of Biofiltration, Biosorption, and Biodegradation Strategies Using Living Fungal Mycelium. *Fermentation.* October 1, **2025**. https://doi.org/10.3390/fermentation11100573.

(53) Sun, J. L.; Zeng, H.; Ni, H. G. Halogenated Polycyclic Aromatic Hydrocarbons in the Environment. *Chemosphere.* Elsevier Ltd **2013**, pp 1751–1759. https://doi.org/10.1016/j.chemosphere.2012.10.094.

(54) Weaver, G. W. Electrophilic Aromatic Substitution. In Organic Reaction Mechanisms 2019: An annual survey covering the literature dated January to December 2019; wiley, **2023**; pp 213–255. https://doi.org/10.1002/9781119608288.ch5b.

(55) Liu, Q.; Xu, X.; Lin, L.; Yang, G.; Wang, D. Occurrence, Health Risk Assessment and Regional Impact of Parent, Halogenated and Oxygenated Polycyclic Aromatic Hydrocarbons in Tap Water. *J. Hazard. Mater*. **2021**, 413. https://doi.org/10.1016/j.jhazmat.2021.125360.

(56) Ohura, T.; Sawada, K. I.; Amagai, T.; Shinomiya, M. Discovery of Novel Halogenated Polycyclic Aromatic Hydrocarbons in Urban Particulate Matters: Occurrence, Photostability, and AhR Activity. *Environ. Sci. Technol.* **2009**, 43 (7), 2269–2275. https://doi.org/10.1021/es803633d.

(57) Berthiaume, A.; Galarneau, E.; Marson, G. Polycyclic Aromatic Compounds (PACs) in the Canadian Environment: Sources and Emissions. *Environmental Pollution*. Elsevier Ltd January 15, **2021**. https://doi.org/10.1016/j.envpol.2020.116008.

(58) Qiao, M.; Qi, W.; Liu, H.; Qu, J. Oxygenated Polycyclic Aromatic Hydrocarbons in the Surface Water Environment: Occurrence, Ecotoxicity, and Sources. *Environment International.* Elsevier Ltd May 1, **2022**. https://doi.org/10.1016/j.envint.2022.107232.

(59) Krzyszczak, A.; Czech, B. Occurrence and Toxicity of Polycyclic Aromatic Hydrocarbons Derivatives in Environmental Matrices. *Science of the Total Environment*. Elsevier B.V. September 20, **2021**. https://doi.org/10.1016/j.scitotenv.2021.147738.

(60) Shin, S. M.; Lee, J. Y.; Shin, H. J.; Kim, Y. P. Seasonal Variation and Source Apportionment of Oxygenated Polycyclic Aromatic Hydrocarbons (OPAHs) and Polycyclic Aromatic Hydrocarbons (PAHs) in PM2.5 in Seoul, Korea. *Atmos. Environ.* **2022**, 272. https://doi.org/10.1016/j.atmosenv.2022.118937.

(61) Blum, P.; Sagner, A.; Tiehm, A.; Martus, P.; Wendel, T.; Grathwohl, P. Importance of Heterocylic Aromatic Compounds in Monitored Natural Attenuation for Coal Tar Contaminated Aquifers: A Review. *Journal of Contaminant Hydrology*. November 1, **2011**, pp 181–194. https://doi.org/10.1016/j.jconhyd.2011.08.004.

(62) Zhihuan, Z.; Fengpeng, H.; Qingwei, B.; Song, L. Vertical Distribution and Environmental Significance of Sulfur and Oxygen Heterocyclic Aromatic Hydrocarbons in Soil Samples Collected from Beijing, China. *Environmental Pollution* **2008**, 153 (2), 457–467. https://doi.org/10.1016/j.envpol.2007.08.004.

(63) Idowu, O.; Carbery, M.; O'Connor, W.; Thavamani, P. Speciation and Source Apportionment of Polycyclic Aromatic Compounds (PACs) in Sediments of the Largest Salt Water Lake of Australia. *Chemosphere* **2020**, 246. https://doi.org/10.1016/j.chemosphere.2019.125779.

(64) Chibwe, L.; Manzano, C. A.; Muir, D.; Atkinson, B.; Kirk, J. L.; Marvin, C. H.; Wang, X.; Teixeira, C.; Shang, D.; Harner, T.; De Silva, A. O. Deposition and Source Identification of Nitrogen Heterocyclic Polycyclic Aromatic Compounds in Snow, Sediment, and Air Samples from the Athabasca Oil Sands Region. *Environ. Sci. Technol.* **2019**, 53 (6), 2981–2989. https://doi.org/10.1021/acs.est.8b06175.

(65) Oz, E. Effects of Smoke Flavoring Using Different Wood Chips and Barbecuing on the Formation of Polycyclic Aromatic Hydrocarbons and Heterocyclic Aromatic Amines in Salmon Fillets. PLoS One **2020**, 15 (1). https://doi.org/10.1371/journal.pone.0227508.

(66) Viegas, O.; Novo, P.; Pinto, E.; Pinho, O.; Ferreira, I. M. P. L. V. O. Effect of Charcoal Types and Grilling Conditions on Formation of Heterocyclic Aromatic Amines (HAs) and Polycyclic Aromatic Hydrocarbons (PAHs) in Grilled Muscle Foods. *Food and Chemical Toxicology* **2012**, 50 (6), 2128–2134. https://doi.org/10.1016/j.fct.2012.03.051.

(67) Adeyeye, S. A. O.; Sivapriya, T.; Sankarganesh, P. Formation and Mitigation of Heterocyclic Amines (HCAs) and Polycyclic Aromatic Hydrocarbons (PAHs) in High Temperature Processed Meat Products: A Review. *Discover Food*. Springer Nature December 1, **2025**. https://doi.org/10.1007/s44187-025-00391-w.

(68) Asif, M.; Wenger, L. M. Heterocyclic Aromatic Hydrocarbon Distributions in Petroleum: A Source Facies Assessment Tool. *Org. Geochem.* **2019**, 137. https://doi.org/10.1016/j.orggeochem.2019.07.005.

(69) Pampanin, D. M.; Sydnes, M. O. Polycyclic Aromatic Hydrocarbons a Constituent of Petroleum: Presence and Influence in the Aquatic Environment. In Hydrocarbon; Kutcherov, V., Kolesnikov, A., Eds.; IntechOpen: Rijeka, **2013**; p Ch. 5. https://doi.org/10.5772/48176.

(70) Mahler, B. J.; Van Metre, P. C.; Foreman, W. T. Concentrations of Polycyclic Aromatic Hydrocarbons (PAHs) and Azaarenes in Runoff from Coal-Tar- and Asphalt-Sealcoated Pavement. *Environmental Pollution* **2014**, 188, 81–87. https://doi.org/10.1016/j.envpol.2014.01.008.

(71) Ahrens, M. J.; Depree, C. V. A Source Mixing Model to Apportion PAHs from Coal Tar and Asphalt Binders in Street Pavements and Urban Aquatic Sediments. *Chemosphere* **2010**, 81 (11), 1526–1535. https://doi.org/10.1016/j.chemosphere.2010.08.030.

(72) Zhang, X.; Bowman, D. T.; Diamond, M. L.; Helm, P.; Jobst, K. J.; Hao, C.; Kleywegt, S.; Zhang, Z. F.; Marvin, C.; Zhang, X. Contribution of Coal Tar Sources to Polycyclic Aromatic Compounds and Associated Ecological Risk in Lake Ontario Sediments: Inference from a Novel Marker. *Environ. Sci. Technol.* **2025**. https://doi.org/10.1021/acs.est.4c10360.

(73) Wang, R.; Sun, R.; Liu, G.; Yousaf, B.; Wu, D.; Chen, J.; Zhang, H. A Review of the Biogeochemical Controls on the Occurrence and Distribution of Polycyclic Aromatic Compounds (PACs) in Coals. *Earth-Science Reviews.* Elsevier B.V. August 1, **2017**, pp 400–418. https://doi.org/10.1016/j.earscirev.2017.06.011.

(74) Ma, Q.; Zhang, X.; Qu, Y. Biodegradation and Biotransformation of Indole: Advances and Perspectives. *Frontiers in Microbiology.* Frontiers Media S.A. November 1, **2018**. https://doi.org/10.3389/fmicb.2018.02625.

(75) Heravi, M. M.; Zadsirjan, V. Prescribed Drugs Containing Nitrogen Heterocycles: An Overview. *RSC Advances*. Royal Society of Chemistry November 23, **2020**, pp 44247–44311. https://doi.org/10.1039/d0ra09198g.

(76) Luongo, G.; Thorsén, G.; Östman, C. Quinolines in Clothing Textiles-A Source of Human Exposure and Wastewater Pollution? *Anal. Bioanal. Chem.* **2014**, 406 (12), 2747–2756. https://doi.org/10.1007/s00216-014-7688-9.

(77) Zhang, B.; Li, X.; Li, B.; Gao, C.; Jiang, Y. Acridine and Its Derivatives: A Patent Review (2009-2013). Expert Opinion on Therapeutic Patents. *Informa Healthcare* **2014**, pp 647–664. https://doi.org/10.1517/13543776.2014.902052.

(78) Luo, Y.; Yue, X.; Wei, P.; Zhou, A.; Kong, X.; Alimzhanova, S. A State-of-the-Art Review of Quinoline Degradation and Technical Bottlenecks. *Science of the Total Environment.* Elsevier B.V. December 10, **2020**. https://doi.org/10.1016/j.scitotenv.2020.141136.

(79) Lellis, B.; Fávaro-Polonio, C. Z.; Pamphile, J. A.; Polonio, J. C. Effects of Textile Dyes on Health and the Environment and Bioremediation Potential of Living Organisms. *Biotechnology Research and Innovation* **2019**, 3 (2), 275–290. https://doi.org/10.1016/j.biori.2019.09.001.

(80) Zou, Y.; Wang, L.; Christensen, E. R. Problems in the Fingerprints Based Polycyclic Aromatic Hydrocarbons Source Apportionment Analysis and a Practical Solution. *Environmental Pollution* **2015**, 205, 394–402. https://doi.org/10.1016/j.envpol.2015.05.029.

(81) Landis, M. S.; Studabaker, W. B.; Patrick Pancras, J.; Graney, J. R.; Puckett, K.; White, E. M.; Edgerton, E. S. Source Apportionment of an Epiphytic Lichen Biomonitor to Elucidate the Sources and Spatial Distribution of Polycyclic Aromatic Hydrocarbons in the Athabasca Oil Sands Region, Alberta, Canada. *Science of the Total Environment* **2019**, 654, 1241–1257. https://doi.org/10.1016/j.scitotenv.2018.11.131.

(82) Xu, Z.; Qian, Y.; Hong, X.; Luo, Z.; Gao, X.; Liang, H. Contamination Characteristics of Polycyclic Aromatic Compounds from Coal Sources in Typical Coal Mining Areas in Huaibei Area, China. *Science of the Total Environment* **2023**, 873. https://doi.org/10.1016/j.scitotenv.2023.162311.

(83) Sofowote, U. M.; Mccarry, B. E.; Marvin, C. H. Source Apportionment of PAH in Hamilton Harbour Suspended Sediments: Comparison of Two Factor Analysis Methods. *Environ. Sci. Technol.* **2008**, 42 (16), 6007–6014. https://doi.org/10.1021/es800219z.

(84) Çelik, G.; Stolte, S.; Müller, S.; Schattenberg, F.; Markiewicz, M. Environmental Persistence Assessment of Heterocyclic Polyaromatic Hydrocarbons – Ultimate and Primary Biodegradability Using Adapted and Non-Adapted Microbial Communities. *J. Hazard. Mater*. **2023**, 460. https://doi.org/10.1016/j.jhazmat.2023.132370.

(85) Ogata, M.; Miyake, Y. Gas Chromatography Combined with Mass Spectrometry for the Identification of Organic Sulfur Compounds in Shellfish and Fish; **1980**; Vol. 18. https://academic.oup.com/chromsci/article/18/11/594/398135.

(86) Novotny, M.; Kump, R.; Merli, F.; Todd, L. J. Capillary Gas Chromatography/Mass Spectrometric Determination of Nitrogen Aromatic Compounds in Complex Mixtures; **1980**; Vol. 52. https://pubs.acs.org/sharingguidelines.

(87) Soursou, V.; Campo, J.; Picó, Y. Revisiting the Analytical Determination of PAHs in Environmental Samples: An Update on Recent Advances. *Trends in Environmental Analytical Chemistry.* Elsevier B.V. March 1, **2023**. https://doi.org/10.1016/j.teac.2023.e00195.

(88) Stout, S. A.; Emsbo-Mattingly, S. D.; Douglas, G. S.; Uhler, A. D.; McCarthy, K. J. Beyond 16 Priority Pollutant PAHs: A Review of PACs Used in Environmental Forensic Chemistry. *Polycycl. Aromat. Compd*. **2015**, 35 (2–4), 285–315. https://doi.org/10.1080/10406638.2014.891144.

(89) Ahad, J. M. E.; Macdonald, R. W.; Parrott, J. L.; Yang, Z.; Zhang, Y.; Siddique, T.; Kuznetsova, A.; Rauert, C.; Galarneau, E.; Studabaker, W. B.; Evans, M.; McMaster, M. E.; Shang,

D. Polycyclic Aromatic Compounds (PACs) in the Canadian Environment: A Review of Sampling Techniques, Strategies and Instrumentation. *Environmental Pollu*tion. Elsevier Ltd November 1, **2020**. https://doi.org/10.1016/j.envpol.2020.114988.

(90) Cargnin, R. S.; Lorenzoni, W. C.; do Nascimento, P. C.; de Carvalho, L. M.; Somavilla, E. A.; Cravo, M. C. C.; do Nascimento, L. A. H. Determination of Polycyclic Aromatic Hydrocarbons (PAHs) and Sulfur, Nitrogen, and Oxygen Analogs in Fumes from Heated Asphalt by High-Performance Liquid Chromatography–Diode Array Detection (HPLC-DAD). *Anal. Lett.* **2024**, 57 (9), 1417–1433. https://doi.org/10.1080/00032719.2023.2250028.

(91) Wnorowski, A.; Aklilu, Y. abeba; Harner, T.; Schuster, J.; Charland, J. P. Polycyclic Aromatic Compounds in Ambient Air in the Surface Minable Area of Athabasca Oil Sands in Alberta (Canada). *Atmos. Environ.* **2021**, 244. https://doi.org/10.1016/j.atmosenv.2020.117897.

(92) Manzano, C. A.; Marvin, C.; Muir, D.; Harner, T.; Martin, J.; Zhang, Y. Heterocyclic Aromatics in Petroleum Coke, Snow, Lake Sediments, and Air Samples from the Athabasca Oil Sands Region. *Environ. Sci. Technol.* **2017**, 51 (10), 5445–5453. https://doi.org/10.1021/acs.est.7b01345.

(93) Marvin, C. H.; Mccarry, B. E.; Villella, J.; Allan, L. M.; Bryant, D. W. Chemical and Biological Profiles of Sediments as Indicators of Sources of Genotoxic Contamination in Hamilton Harbour. Part I: Analysis of Polycyclic Aromatic Hydrocarbons and Thia-Arene Compounds.

(94) Bahia, P. V. B.; Nascimento, M. M.; Hatje, V.; de Andrade, J. B.; Machado, M. E. Microscale Extraction Combined with Gas Chromatography/Mass Spectrometry for the Simultaneous Determination of Polycyclic Aromatic Hydrocarbons and Polycyclic Aromatic Sulfur Heterocycles in Marine Sediments. *J. Chromatogr. A* **2021**, 1653. https://doi.org/10.1016/j.chroma.2021.462414.

(95) Witter, A. E.; Nguyen, M. H. Determination of Oxygen, Nitrogen, and Sulfur-Containing Polycyclic Aromatic Hydrocarbons (PAHs) in Urban Stream Sediments. *Environmental Pollution* **2016**, 209, 186–196. https://doi.org/10.1016/j.envpol.2015.10.037.

(96) Brack, W.; Schirmer, K. Effect-Directed Identification of Oxygen and Sulfur Heterocycles as Major Polycyclic Aromatic Cytochrome P4501A-Inducers in a Contaminated Sediment. *Environ. Sci. Technol*. **2003**, 37 (14), 3062–3070. https://doi.org/10.1021/es020248j.

(97) Musa Bandowe, B. A.; Wei, C.; Han, Y.; Cao, J.; Zhan, C.; Wilcke, W. Polycyclic Aromatic Compounds (PAHs, Oxygenated PAHs, Nitrated PAHs and Azaarenes) in Soils from China and Their Relationship with Geographic Location, Land Use and Soil Carbon Fractions. *Science of the Total Environment* **2019**, 690, 1268–1276. https://doi.org/10.1016/j.scitotenv.2019.07.022.

(98) Dreij, K.; Lundin, L.; Le Bihanic, F.; Lundstedt, S. Polycyclic Aromatic Compounds in Urban Soils of Stockholm City: Occurrence, Sources and Human Health Risk Assessment. *Environ. Res.* **2020**, 182. https://doi.org/10.1016/j.envres.2019.108989.

(99) Ai, Q.; Gao, L.; Huang, D.; Yang, J.; Fu, Q.; Zheng, X.; Liu, Y.; Qiao, L.; Weng, J.; Zheng, M. Non-Target and Target Analysis to Identify and Characterize Thiophenes in Soil from an Abandoned Coking Plant. *J. Hazard. Mater.* **2023**, 460. https://doi.org/10.1016/j.jhazmat.2023.132444.

(100) Kellems, B.; Fabian, K.; Eng, P.; Orchard, B.; Parmelee, R.; Bonkoski, B.; Spadaro, P.; Beaver, J.; Fitzgerald, W. P.; Santiago, R.; Rozenberg, W. Randle Reef Sediment Remediation Project; **2013**.

(101) Burniston, D. A.; Jia, J.; Charlton, M. N.; Thiessen, L.; McCarry, B. E.; Marvin, C. H. Trends in Hamilton Harbour Suspended Sediment Quality. *Aquat. Ecosyst. Health Manag*. **2016**, 19 (2), 141–149. https://doi.org/10.1080/14634988.2016.1172906.

(102) Chibwe, L.; Roberts, S.; Shang, D.; Yang, F.; Manzano, C. A.; Wang, X.; Kirk, J. L.; Muir, D. C. G. A One-Century Sedimentary Record of N- and S-Polycyclic Aromatic Compounds in the Athabasca Oil Sands Region in Canada. *Chemosphere* **2020**, 260. https://doi.org/10.1016/j.chemosphere.2020.127641.

(103) Satilmis, I.; Schrader, W. Studying the Fenton Treatment of Polycyclic Aromatic Compounds in a Highly Contaminated Soil with Different Modifiers by High Resolution Mass Spectrometry. *Journal of Hazardous Materials Advances* **2022**, 8. https://doi.org/10.1016/j.hazadv.2022.100200.

(104) Golzadeh, N.; Barst, B. D.; Baker, J. M.; Auger, J. C.; McKinney, M. A. Alkylated Polycyclic Aromatic Hydrocarbons Are the Largest Contributor to Polycyclic Aromatic Compound Concentrations in Traditional Foods of the Bigstone Cree Nation in Alberta, Canada. *Environmental Pollution* **2021**, 275. https://doi.org/10.1016/j.envpol.2021.116625.

(105) Provencher, J. F.; Thomas, P. J.; Braune, B. M.; Pauli, B.; Tomy, G.; Idowu, I.; O'Hara, P.; Mallory, M. L. Decadal Differences in Polycyclic Aromatic Compound (PAC) Concentrations in Two Seabird Species in Arctic Canada. *Science of the Total Environment* **2022**, 826. https://doi.org/10.1016/j.scitotenv.2022.154088.

(106) Fernie, K. J.; Marteinson, S. C.; Chen, D.; Eng, A.; Harner, T.; Smits, J. E. G.; Soos, C. Elevated Exposure, Uptake and Accumulation of Polycyclic Aromatic Hydrocarbons by Nestling Tree Swallows (Tachycineta Bicolor) through Multiple Exposure Routes in Active Mining-Related Areas of the Athabasca Oil Sands Region. *Science of the Total Environment* **2018**, 624, 250–261. https://doi.org/10.1016/j.scitotenv.2017.12.123.

(107) Li, A.; Zhou, S.; He, H.; Guo, J.; Rockne, K. J.; Sturchio, N. C.; Giesy, J. P. Polyhalogenated Carbazoles in Sediments of Lower Laurentian Great Lakes and Regional Perspectives. ACS *ES and T Water* **2022**, 2 (9), 1544–1554. https://doi.org/10.1021/acsestwater.2c00191.

(108) Guo, J.; Li, Z.; Ranasinghe, P.; Bonina, S.; Hosseini, S.; Corcoran, M. B.; Smalley, C.; Rockne, K. J.; Sturchio, N. C.; Giesy, J. P.; Li, A. Spatial and Temporal Trends of Polyhalogenated

Carbazoles in Sediments of Upper Great Lakes: Insights into Their Origin. *Environ. Sci. Technol.* **2017**, 51 (1), 89–97. https://doi.org/10.1021/acs.est.6b06128.

(109) Norstrom, R. J. Polychlorinated Dibenzo-p-Dioxins and Dibenzofurans in the Great Lakes. Handbook of Environmental Chemistry, Volume 5: *Water Pollution*. **2006**, pp 71–150. https://doi.org/10.1007/698_5_040.

(110) Walker, M. K. , and R. E. P. Potencies of Polychlorinated Dibenzo-p-Dioxin, Dibenzofuran, and Biphenyl Congeners, Relative to 2,3,7,8-Tetrachlorodibenzo-p-Dioxin, for Producing Early Life Stage Mortality in Rainbow Trout (Oncorhynchus Mykiss). *Aquat. Toxicol*. **1991**, 20 ((3/4)), 219–238.

(111) Idowu, O.; Tran, T. K. A.; Baker, P.; Farrel, H.; Zammit, A.; Semple, K. T.; O'Connor, W.; Thavamani, P. Bioavailability of Polycyclic Aromatic Compounds (PACs) to the Sydney Rock Oyster (Saccostrea Glomerata) from Sediment Matrices of an Economically Important Australian Estuary. *Science of the Total Environment* **2020**, 736. https://doi.org/10.1016/j.scitotenv.2020.139574.

(112) Perono, G. A.; Tomy, T.; Loudon, K.; Jamshed, L.; Garlisi, B.; Lauks, S.; Lockington, C.; Ruan, C.; Tomy, G. T.; Petrik, J. J.; Thomas, P. J.; Holloway, A. C. Sulfur-Containing Heterocyclic Aromatic Hydrocarbons Alter Estrogen Metabolism and Cause DNA Damage and Apoptosis in Granulosa Cells. *Int. J. Mol. Sci.* **2025**, 26 (16). https://doi.org/10.3390/ijms26168004.

(113) Muscalu, A. M.; Górecki, T. Comprehensive Two-Dimensional Gas Chromatography in Environmental Analysis. *TrAC - Trends in Analytical Chemistry.* Elsevier B.V. September 1, **2018**, pp 225–245. https://doi.org/10.1016/j.trac.2018.07.001.

(114) Veenaas, C.; Bignert, A.; Liljelind, P.; Haglund, P. Nontarget Screening and Time-Trend Analysis of Sewage Sludge Contaminants via Two-Dimensional Gas Chromatography-High Resolution Mass Spectrometry. *Environ. Sci. Technol.* **2018**, 52 (14), 7813–7822. https://doi.org/10.1021/acs.est.8b01126.

(115) Møller, M. T.; Birch, H.; Sjøholm, K. K.; Papazian, S.; Wennberg, A. C.; Bonnefille, B.; Kronsbein, P. M.; Kelland, M. A.; Martin, J. W.; Mayer, P. Persistence-Directed Testing of Chemicals Discharged from Offshore Oil Platforms Combined with Nontargeted Analysis. *Environ. Sci. Technol.* **2025**, 59 (44), 24000–24011. https://doi.org/10.1021/acs.est.5c08802.

(116) cker, G.; Nilsson, U.; Colmsjo, A.; Ostman, C. Determination of Polycyclic Aromatic Sulfur Heterocyclic Compounds in Airborne Particulate by Gas Chromatography with Atomic Emission and Mass Spectrometric Detection. *J. Chromatogr. A* **1998**, *826*, 57–66.

(117) Sebastiani, B.; Mariucci, S.; Palmieri, N. Foraging Honeybees (Apis Mellifera Ligustica) as Biocenosis Monitors of Pollution in Areas Affected by Cement Industry Emissions. *Environments* - MDPI 2023, 10 (2). https://doi.org/10.3390/environments10020034.

(118) Elliott, S. M.; Brigham, M. E.; Lee, K. E.; Banda, J. A.; Choy, S. J.; Gefell, D. J.; Minarik, T. A.; Moore, J. N.; Jorgenson, Z. G. Contaminants of Emerging Concern in Tributaries to the

Laurentian Great Lakes: I. Patterns of Occurrence. *PLoS One* **2017**, 12 (9). https://doi.org/10.1371/journal.pone.0182868.

(119) Witter, A. E.; Nguyen, M. H. Determination of Oxygen, Nitrogen, and Sulfur-Containing Polycyclic Aromatic Hydrocarbons (PAHs) in Urban Stream Sediments. *Environmental Pollution* **2016**, 209, 186–196. https://doi.org/10.1016/j.envpol.2015.10.037.

(120) Song, Q.; Xiao, S.; Zeng, X.; Zhang, B.; Zhu, Z.; Liang, Y.; Yu, Z. Presence of Polycyclic Aromatic Compounds in River Sediment and Surrounding Soil: Possible Impact from Shale Gas Wastewater. *Science of the Total Environment* **2024**, 953. https://doi.org/10.1016/j.scitotenv.2024.176186.

(121) Titaley, I. A.; Trine, L. S. D.; Wang, T.; Duberg, D.; Davis, E. L.; Engwall, M.; Massey Simonich, S. L.; Larsson, M. Extensive Chemical and Bioassay Analysis of Polycyclic Aromatic Compounds in a Creosote-Contaminated Superfund Soil Following Steam Enhanced Extraction. *Environmental Pollution* **2022**, 312. https://doi.org/10.1016/j.envpol.2022.120014.

(122) Idowu, O.; Semple, K. T.; Ramadass, K.; O'Connor, W.; Hansbro, P.; Thavamani, P. Analysis of Polycyclic Aromatic Hydrocarbons (PAHs) and Their Polar Derivatives in Soils of an Industrial Heritage City of Australia. *Science of the Total Environment* **2020**, 699. https://doi.org/10.1016/j.scitotenv.2019.134303.

(123) Lundstedt, S.; Bandowe, B. A. M.; Wilcke, W.; Boll, E.; Christensen, J. H.; Vila, J.; Grifoll, M.; Faure, P.; Biache, C.; Lorgeoux, C.; Larsson, M.; Frech Irgum, K.; Ivarsson, P.; Ricci, M. First Intercomparison Study on the Analysis of Oxygenated Polycyclic Aromatic Hydrocarbons (Oxy-PAHs) and Nitrogen Heterocyclic Polycyclic Aromatic Compounds (N-PACs) in Contaminated Soil. *TrAC - Trends in Analytical Chemistry*. Elsevier B.V. **2014**, pp 83–92. https://doi.org/10.1016/j.trac.2014.01.007.

(124) Schlanges, I.; Meyer, D.; Palm, W. U.; Ruck, W. Identification, Quantification and Distribution of PAC-Metabolites, Heterocyclic PAC and Substituted PAC in Groundwater Samples of Tar-Contaminated Sites from Germany. *Polycyclic Aromatic Compounds;* Taylor and Francis Inc., **2008**; Vol. 28, pp 320–338. https://doi.org/10.1080/10406630802377807.

(125) Cook, J. W. .; Hewett, C. L.; Hieger, I. Isolation of a Cancer-Producing Hydrocarbon from Coal Tar. Part I, II, and III. J. Chem. Soc. **1933**.

(126) Chemical Agents and Related Occupations Volume 100 F A Review of Human Carcinogens; **2012**.

(127) Billiard, S. M. S. M.; Timme-Laragy, A. R.; Wassenberg, D. M.; Cockman, C.; Di Giulio, R. T. The Role of the Aryl Hydrocarbon Receptor Pathway in Mediating Synergistic Developmental Toxicity of Polycyclic Aromatic Hydrocarbons to Zebrafish. *Toxicological Sciences* **2006**, 92 (2), 526–536. https://doi.org/10.1093/toxsci/kfl011.

(128) Kim, D.; Guengerich, F. P. Cytochrome P450 Activation of Arylamines and Heterocyclic Amines. *Annual Review of Pharmacology and Toxicology.* **2005**, pp 27–49. https://doi.org/10.1146/annurev.pharmtox.45.120403.100010.

(129) Xue, W.; Warshawsky, D. Metabolic Activation of Polycyclic and Heterocyclic Aromatic Hydrocarbons and DNA Damage: A Review. *Toxicol. Appl. Pharmacol*. **2005**, 206 (1), 73–93. https://doi.org/10.1016/j.taap.2004.11.006.

(130) Liebler, D. C.; Guengerich, F. P. Elucidating Mechanisms of Drug-Induced Toxicity. *Nature Reviews Drug Discovery.* May **2005**, pp 410–420. https://doi.org/10.1038/nrd1720.

(131) Bellamri, M.; Walmsley, S. J.; Turesky, R. J. Metabolism and Biomarkers of Heterocyclic Aromatic Amines in Humans. Genes and Environment. *BioMed Central* Ltd December 1, **2021**. https://doi.org/10.1186/s41021-021-00200-7.

(132) Balaguer, P.; Delfosse, V.; Bourguet, W. Mechanisms of Endocrine Disruption through Nuclear Receptors and Related Pathways. Current Opinion in Endocrine and Metabolic Research. Elsevier Ltd August 1, **2019**, pp 1–8. https://doi.org/10.1016/j.coemr.2019.04.008.

(133) Vrzalová, A.; Vrzal, R.; Nádvorník, P.; Šebela, M.; Dvořák, Z. Modulation of Aryl Hydrocarbon Receptor Activity by Halogenated Indoles. *Bioorg. Med. Chem*. **2024**, 114. https://doi.org/10.1016/j.bmc.2024.117964.

(134) Vondra´čvondra´, J.; Ek, V.; Kateř, §; Chramostova´, I.; Chramostova´, C.; Pli´š, M.; Pli´š, P.; Kova´kova´, K.; Ludě, ‡; Bla´ha, K.; Bla´ha, B.; Brack, W. Induction of aryl hydrocarbon receptor-mediated and estrogen receptor-mediated activities, and modulation of cell proliferation by dinaphthofurans; **2004**; Vol. 23. https://academic.oup.com/etc/article/23/9/2214/7775272.

(135) Baldwin, A. K.; Corsi, S. R.; Stefaniak, O. M.; Loken, L. C.; Villeneuve, D. L.; Ankley, G. T.; Blackwell, B. R.; Lenaker, P. L.; Nott, M. A.; Mills, M. A. Risk-Based Prioritization of Organic Chemicals and Locations of Ecological Concern in Sediment From Great Lakes Tributaries. *Environ. Toxicol. Chem*. **2022**, 41 (4), 1016–1041. https://doi.org/10.1002/etc.5286.

(136) Barzegar, F.; Kamankesh, M.; Mohammadi, A. Heterocyclic Aromatic Amines in Cooked Food: A Review on Formation, Health Risk-Toxicology and Their Analytical Techniques. *Food Chem.* **2019**, 280, 240–254. https://doi.org/10.1016/j.foodchem.2018.12.058.

(137) Bleeker, E. A. J.; Pieters, B. J.; Wiegman, S.; Kraak, M. H. S. Comparative (Photoenhanced) Toxicity of Homocyclic and Heterocyclic PACs. *Polycyclic Aromatic Compounds*; Taylor and Francis Ltd., **2002**; Vol. 22, pp 601–610. https://doi.org/10.1080/10406630290103771.

(138) Chlebowski, A. C.; Garcia, G. R.; La Du, J. K.; Bisson, W. H.; Truong, L.; Simonich, S. L. M.; Tanguay, R. L. Mechanistic Investigations into the Developmental Toxicity of Nitrated and Heterocyclic PAHs. *Toxicological Sciences* **2017**, 157 (1), 246–259. https://doi.org/10.1093/toxsci/kfx035.

(139) Le Bihanic, F.; Morin, B.; Cousin, X.; Le Menach, K.; Budzinski, H.; Cachot, J. Developmental Toxicity of PAH Mixtures in Fish Early Life Stages. Part I: Adverse Effects in Rainbow Trout. *Environmental Science and Pollution Research* **2014**, 21 (24), 13720–13731. https://doi.org/10.1007/s11356-014-2804-0.

(140) León Paumen, M.; de Voogt, P.; van Gestel, C. A. M.; Kraak, M. H. S. Comparative Chronic Toxicity of Homo- and Heterocyclic Aromatic Compounds to Benthic and Terrestrial Invertebrates: Generalizations and Exceptions. *Science of the Total Environment* **2009**, 407 (16), 4605–4609. https://doi.org/10.1016/j.scitotenv.2009.04.039.

(141) Peddinghaus, S.; Brinkmann, M.; Bluhm, K.; Sagner, A.; Hinger, G.; Braunbeck, T.; Eisenträger, A.; Tiehm, A.; Hollert, H.; Keiter, S. H. Quantitative Assessment of the Embryotoxic Potential of NSO-Heterocyclic Compounds Using Zebrafish (Danio Rerio). *Reproductive Toxicology* **2012**, 33 (2), 224–232. https://doi.org/10.1016/j.reprotox.2011.12.005.

(142) Sovadinova´, I.; Sovadinova´, S.; Bla´ha, L. K.; Bla´ha, B.; Jaroslav, †; Ek, J.; Kla´, K.; Hilscherova´, K.; Hilscherova´, H.; Giesy, J. P.; Jones, P. D.; Holoubek, I. Cytotoxicity and aryl hydrocarbon receptor-mediated activity of n-heterocyclic polycyclic aromatic hydrocarbons: structure-activity relationships; **2006**; Vol. 25. https://academic.oup.com/etc/article/25/5/1291/7771412.

(143) Wassenberg, D. M.; Nerlinger, A. L.; Battle, L. P.; Di Giulio, R. T. Effects of the polycyclic aromatic hydrocarbon heterocycles, carbazole and dibenzothiophene, on in vivo and in vitro cyp1a activity and polycyclic aromatic hydrocarbon-derived embryonic deformities; **2005**; Vol. 24. https://academic.oup.com/etc/article/24/10/2526/7770615.

# CHAPTER 2 METHODOLOGY

Analytical Chemistry is based on generating reliable and reproducible chemical measurements. The integrity of any analytical method is determined by the robustness of the entire workflow from sample collection to data interpretation. This workflow generally consists of proper sample preparation and purification techniques, standard quality assurance and quality control (QA/QC) guidelines, optimized data acquisition and data processing methods, and data interpretation with appropriate statistical methods. When all these parameters are systematically integrated, they collectively produce high-quality data that can form a reliable foundation for informed decision-making, including environmental monitoring, risk assessment, and the development of regulatory policies. This chapter presents a comprehensive description of all the steps used in my thesis.

## 2.1 SAMPLE PREPARATION

Organic pollutants such as PACs are typically present at trace levels in environmental matrices. The inherently complex nature of environmental samples requires meticulous attention to sample preparation and purification steps to ensure selective detection of analytes of interest while minimizing matrix effects. In recent years, sample preparation for complex matrices such as sediments, fish muscle, and bird eggs has incorporated advanced extraction and cleanup techniques to meet these requirements and improve analytical performance.

### 2.1.1 Sample collection and pre-preparation

#### 2.1.1.1 Sediment samples

Sixty-nine sediment samples that contained both suspended and bottom sediments were analyzed in this study. These samples were collected from three sampling areas in Lake Ontario, North America, between 2014 and 2023. The three areas can be named as Hamilton Harbour (HH) area,

Toronto Harbour (TH) area, and nearby river mouth areas, which are named as Non-Harbour (NH) area. The detailed information on sampling sites and sample collection is published in Idowu et al.[1] and Zhang et al.[2]

Hamilton Harbour (also known as Burlington Harbour; 43°17′ N, 79°50′ W; ~2150 ha; mean depth 13 m), which is one of the major shipping hubs, is located at the western end of Lake Ontario, bordered by the cities of Hamilton and Burlington, Ontario.[3] The area is characterized by a mix of heavy and light industrial activities, including steel manufacturing. The Harbour serves as both a source and a recipient of cooling water, stormwater, and treated municipal wastewater effluent from surrounding urban areas and the watershed, which has diverse land uses. All the suspended sediment samples were collected from HH using sediment traps. Sampling was conducted seasonally (May to November) across multiple locations with varying environmental conditions, including hydrodynamics and pollution sources. Collected samples were transported to the laboratory, where overlying water was carefully removed, and the retained particulate matter was transferred to containers and frozen until further analysis.

Toronto Harbour, also referred to as Toronto Bay (43°38′ N, 79°22′ W), and Humber Bay (43°38′ N, 79°27′ W) are located along the northern shore of Lake Ontario adjacent to the City of Toronto. These water bodies receive stormwater inputs from sewer systems, inflow from the Don River, and are influenced by a small municipal wastewater treatment plant. Toronto Harbour accommodates shipping activities, passenger ferries, and various recreational and commercial vessels. Humber Bay, which receives runoff from the Humber River and Mimico Creek, also contains a major wastewater treatment plant diffuser.[1]

A total of eighteen bottom sediment samples were collected in 2014 and 2018 from all three sampling areas. Sampling locations were selected to represent key environmental features such as river inputs, wastewater treatment plant (WWTP) discharges, combined sewer outfalls (CSOs), central harbour zones, and offshore locations.

Sediment samples were collected using appropriate grab sampling techniques depending on location, including a petite ponar or stainless-steel spoon for rivers and creeks, and a Shipek grab sampler for lake sites. The top 2–3 cm of at least three grabs was homogenized to ensure representativeness. Samples were processed in solvent-rinsed equipment, transferred into pre-cleaned amber glass jars, and stored frozen until analysis. These sediment samples were used in the analysis I have described in Chapters 4 and 8.

#### 2.1.1.2 Fish samples

Ten different fish species were collected from the three basins of Lake Erie (42.2°N 81.2°W) between May and October 2019 as part of the Lake Erie CSMI program (Forage Task Group, 2023). The full sampling procedure was previously published in Heuvel et al.[4,5] The target sample size was a minimum of five individuals per basin, with an optimal goal of ten. Fish were euthanized according to University of Windsor animal care protocol under CCAC guidelines and in accordance with standard operating procedures established by the participating agencies (the Ontario Ministry of Natural Resources, the Ohio Department of Natural Resources, and the U.S. Geological Survey) and relevant animal care guidelines. Specimens were then placed on ice and transported to the Great Lakes Institute for Environmental Research at the University of Windsor, Ontario, Canada.

During dissection, the total length and wet weight of each fish were recorded. A 1–2 g sample of skinless muscle tissue was excised from the region below the dorsal fin and above the lateral line for stable isotope analysis. All samples were stored at −20 °C prior to analysis. These fish samples were used in the studies I have described in Chapters 3, 4, and 8.

### 2.1.2 Extraction and purification

The choice of extraction methods, cleanup strategies, and concentration steps was carefully tailored to the p-chem properties of the analytes and the sample matrix. The extraction and purification methods were first developed and validated by Xia et al. to capture non-polar PACs[6] In 2021, I modified these methods to co-extract more polar HPACs with all other PAC classes. To enhance the selectivity of target PACs over matrix interferences, and to ensure optimal performance of the GC-MS instrumentation, I implemented additional purification steps. These procedures also contributed to improved detector sensitivity and overall analytical reliability. Extraction and purification procedures were adapted to the sample matrices as follows.

#### 2.1.2.1 Sediment samples

Sediment samples, collected from Lake Ontario, were received from the Water Science and Technology Directorate, Environment and Climate Change Canada, Burlington, Ontario, CA, and the Ministry of Environment, Conservation and Parks, Toronto, Ontario, CA. Prior to the extraction, the wet sediments were freeze-dried for 48 hours using a LABCONCO freeze drier. Then, 1g of dry sediment samples was extracted using the one-step accelerated solvent extraction (ASE) method developed by Xia et al.[6] The extracting solvent is changed from 100% DCM to a 70:30 DCM: ethyl acetate mixture to improve the extraction efficiency of HPACs.[7]

Accelerated solvent extraction is a type of pressurized solvent extraction (PLE) that has been used to extract PACs since the 1990s.[8] The technique uses high temperature and pressure conditions to exhaustively extract tightly bound non-polar organic compounds such as PACs and PCBs from sediment samples. Compared to traditional extraction techniques like Soxhlet extraction, ASE is more efficient and greener with the use of low solvent volumes.[9–11]

In the one-step ASE method for sediments, the extraction and purification steps are combined to increase the sample preparation efficiency. A mixture of 1 g of reduced copper, 4.5 g of 5% deactivated alumina, and 5 g of silica gel was put into 34 mL ASE cells with glass fiber filters at the bottom. Samples were accurately weighed, homogenized with 1.5 g of diatomaceous earth (DE) as a dispersant, and transferred into the ASE cells. Each sample was subsequently spiked with a recovery internal standard (RIS) to account for potential analyte losses during sample preparation. Finally, Ottawa sand was added to fill any remaining void space within the cell.[6]

Freshly prepared reduced copper was used to remove inorganic sulfur ($S_8$) from the sediments, as it has the potential to co-extract and co-elute with PACs.[12] Copper powder purchased from Fisher Chemicals (Ottawa, Canada) was reduced using 5% nitric acid solution and washed with water and acetone prior to drying. This reduced copper forms copper sulfide with $S_8$ and remains in the ASE cell. Silica ($SiO_2$) and alumina ($Al_2O_3$) were used to retain polar basic and acidic compounds, respectively, in the solid mixture. These can bind to polar compounds primarily through surface hydroxyl groups (-OH) that act as acidic or basic sites, enabling acid-base interactions, hydrogen bonding, and electrostatic repulsion or attraction.[13,14]

The ASE extraction conditions were set as follows: system pressure of 1500 psi, oven temperature of 120 °C, a static time of 15 minutes with two static cycles, using the 1:1 DCM: ethyl acetate

mixture as the extraction solvent, a flush volume of 60%, and a nitrogen purge for 80 seconds at 150 psi. Following extraction, the extracts were dried by passing them through a glass funnel containing sodium sulfate and subsequently concentrated to 100 µL under a gentle stream of ultra-high purity (UHP) nitrogen. The extracts were then transferred to GC vials, spiked with 10.0 µL of a 10 ng/µL instrument performance internal standard (IPIS), and diluted to a final volume of 1 mL with hexane. The sediment extraction and purification method was validated according to the Eurachem guidelines and validation procedure, and the performance characteristics are described in Chapter 3.

#### 2.1.2.2 Coal tar and non-coal tar based reference materials

Fourteen (14) reference materials including tire rubber (> 25 years old; > 5 years old), two CT based reference materials (polymeric CT; CT emulsion), two asphalt-based sealers (NCT, reference material), raw coal tar, tar from tar & chip roadway, creosote-treated railway tie pieces, aged asphalt shingle pieces, airport-grade driveway resurfacer-repair non-fibrous roof coating, commercial roof tar and crack filler-liquid were either purchased or obtained from the environment to identify the potential sources of PACs in Lake Ontario sediments (*see* Chapter 4). The extraction and purification procedure used for these reference materials was previously published in Vitharana et al.[15]

Reference materials were extracted using a custom-made 50 mL stainless steel extraction cell designed for use with a Precellys Evolution Homogenizer (Bertin Technologies, Montigny-le-Bretonneux, France). Approximately 0.1 g of each of the 14 reference materials was weighed into separate extraction cells, spiked with 10 µL of a 10 ng/µL RIS, and extracted individually with 24 mL of a 1:1 (v/v) DCM: ethyl acetate mixture at 1000 g for 5 minutes.

Extracts were cleaned twice using 20 g of silica gel in a dispersive solid-phase extraction setup. Prior to use, the silica gel was pre-cleaned with 20 mL of DCM followed by 20 mL of ethyl acetate, air-dried, and then packed into a 75 mL polypropylene cartridge using DCM. Samples were loaded onto the cartridge, and sequential elution was performed with 30 mL of DCM (Fraction 1), followed by 20 mL of ethyl acetate (Fraction 2). The collected fractions were combined, concentrated, and further purified by gel permeation chromatography (GPC) using a 50:50 (v/v) DCM: hexane mixture.

Following GPC, the extracts were reduced to approximately 2 mL in 250 mL flat-bottom flasks using a 70:30 (v/v) hexane: DCM mixture. A cleanup sorbent mixture consisting of 4 g silica gel, 0.5 g sodium sulfate, and 1 g of 5% deactivated alumina was then added to each flask. The flasks were allowed to stand for 30 minutes with intermittent swirling to facilitate further purification. Subsequently, the extracts were transferred to 60 mL ASE tubes using one 5 mL rinse with 70:30 hexane: DCM, followed by two 10 mL ethyl acetate rinses. Final volumes were reduced to 1 mL and analyzed by GC/MS-MS.

#### 2.1.2.3 Fish samples

Around 250 fish muscle tissue samples collected from Lake Erie were received from the Great Lakes Institute for Environmental Research (GLIER), University of Windsor, Windsor, Ontario, CA. Approximately 5 g of wet fish muscle tissue was freeze-dried for 48 hours using a LABCONCO freeze dryer. The dried samples were then homogenized with 3.5 g of baked diatomaceous earth (DE) dispersant and loaded into accelerated solvent extraction (ASE) cells fitted with a glass fiber filter at the bottom. Each cell was spiked with 10.0 µL of a 1 ng/µL recovery

internal standard (RIS) mixture and packed with baked Ottawa sand to fill any remaining void volume.

Extraction was carried out using a Dionex ASE 350 system (Thermo Fisher Scientific, Mississauga, ON) using the same method as the sediment samples. Following extraction, samples were dried by adding anhydrous sodium sulfate to the extracts and filtering through glass wool. The extract volume was reduced to 5 mL, and a 0.5 mL aliquot was removed for gravimetric lipid determination. The remaining extract was subjected to GPC using a J2 Scientific AccuPrepMPS system (Columbia, Missouri, USA) equipped with a column packed with 60 g of S-X3 Bio-Beads. A 50:50 (v/v) DCM: hexane mixture was used as the mobile phase at a pressure of 8 psi and a flow rate of 5 mL/min.

Gel permeation chromatography is a type of size-exclusion chromatography used to remove larger molecules, such as lipids and pigments, from sample extracts. The micro-pores in the beads can separate smaller molecules from larger ones by increasing their elution time as the flow path length increases.[16] GPC fraction collection was optimized by spiking native analytical standards into 1 g of cod liver oil and processing under identical conditions.[17] The initial 120 mL fraction, containing lipids and other co-extracted matrix components, was discarded, while the fraction between 120 mL and 250 mL was collected for further analysis. This fraction was then concentrated to approximately 2 mL using a rotary evaporator.

Further cleanup was performed using a pre-cleaned sorbent mixture consisting of 4 g of silica gel, 1 g of 5% deactivated alumina, and 0.5 g of sodium sulfate. The sorbents were prewashed with 10 mL of 70:30 (v/v) DCM: hexane, followed by 10 mL of ethyl acetate, and then transferred into the sample-containing flask using two 10 mL rinses of 70:30 (v/v) DCM: hexane. The mixture was

allowed to stand for 30 minutes with intermittent swirling. The extracts were then transferred to 60 mL test tubes through glass wool, using two 5 mL rinses of 70:30 (v/v) DCM: hexane and a final 10 mL ethyl acetate rinse. The volume was reduced to 100 µL under a gentle stream of ultra-high purity (UHP) nitrogen, spiked with 10.0 µL of a 1 ng/µL instrument performance internal standard (IPIS; anthracene-d10), and prepared for GC-MS/MS analysis. The method was validated according to the Eurachem guidelines and validation procedure, and the performance characteristics are described in Chapter 5.

## 2.2 DATA ACQUISITION AND QUANTIFICATION

Optimization of data acquisition parameters further strengthens the reliability of analytical methods. Instrumental conditions such as resolution, sensitivity, mass accuracy, and signal-to-noise (S/N) ratios should be carefully selected to capture meaningful chemical information, avoiding false positives and false negatives. Several analytical techniques are available for analyzing PACs in environmental samples. Chromatographic methods such as high-performance liquid chromatography (HPLC) and gas chromatography (GC), coupled with detectors including diode array detectors (DAD),[18] flame ionization detection (FID),[19] fluorescence spectroscopy,[20] UV–visible absorption spectroscopy,[21] and mass spectrometry (MS),[22,23] are among the most commonly used approaches. Among these, GC coupled with mass spectrometry has been employed in most studies for several key reasons.

Polycyclic aromatic compounds are generally non-polar, hydrophobic, semi-volatile, and thermally stable, making them more amenable to separation by GC than by HPLC.[24] For compound identification, mass spectrometry offers significant advantages over other detection techniques due to its ability to characterize a wide range of PACs with high specificity.[25,26] In

addition, the aromatic structure of PACs often allows the molecular ion to be preserved under both electron ionization (EI) and chemical ionization (CI), facilitating more accurate compound identification.

In my work, two primary analytical platforms were employed to detect and quantify PACs in environmental samples. Targeted analysis of approximately 120 PACs was performed using gas chromatography coupled with tandem mass spectrometry (GC–MS/MS). In parallel, comprehensive two-dimensional gas chromatography coupled with high-resolution time-of-flight mass spectrometry (GC×GC–HRTOF-MS) was used for non-targeted analysis. The following sections describe the instrumentation and method parameters for these analytical techniques.

### 2.2.1 Gas chromatography

In 1941, two British scientists, Archer J. P. Martin and Richard L. M. Synge, introduced the theoretical foundation of gas chromatography through the concept of partition chromatography.[27] In their work, they proposed that using a gaseous mobile phase, rather than a liquid, could enable the separation of volatile compounds by their differential partitioning between a stationary liquid phase and a flowing gas phase. For this pioneering contribution, they were awarded the Nobel Prize in Chemistry in 1952. The technique was first experimentally demonstrated in 1952 by Anthony T. James and Archer J. P. Martin and was initially applied to the separation of volatile fatty acids.[28,29]

Since its development, GC has become one of the primary techniques for the separation and identification of organic pollutants. In GC analysis, samples are introduced into a heated inlet where they are rapidly vaporized. The volatilized analytes are then carried by an inert gas through a capillary column coated with a liquid stationary phase. The GC column is in the GC oven, where

the temperature changes according to a given program, and compounds are preliminarily separated based on their boiling points. Then, the separated compounds are directed to a detector, in this case a mass spectrometer.

### 2.2.2 Mass spectrometry

Mass spectrometry (MS) originated in the early 20th century with studies by J. J. Thomson on ionized particles and their behavior in electric and magnetic fields. In 1912, he demonstrated that ions could be separated based on their mass to charge (m/z) ratios and identified neon isotopes, establishing the fundamental principle of MS.[30] Since then, numerous scientific advancements contributed to the highly advanced modern mass spectrometric techniques that are able to identify compounds with parts per million (ppm) mass accuracy levels.[31–33]

A mass spectrometer consists of three main components: an ion source, a mass analyzer, and a detector. As the names suggest, the ion source produces ions from the injected chemicals, and the mass analyzer sorts them by m/z ratio, which are then identified by the detector at the end.

The most common ion sources used in MS are electron ionization (EI) and chemical ionization (CI).[31] Electron ionization is widely used in GC-MS for volatile and semi-volatile organic compounds like PACs. Here, electrons at 70 eV collide with analytes, producing molecular ions and fragments for structural identification. This produces highly reproducible spectra for library matching.[34] A schematic diagram of an EI is depicted in Figure 2.1. Chemical ionization (CI) is a soft ionization technique compared to EI. The fragmentation achieved via reagent gas reactions (e.g., methane PCI or NCI for electron-capturing compounds) improves sensitivity/selectivity for isomers in environmental samples.[35]

Electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI) are commonly used in LC-MS for polar and thermally unstable compounds, such as pharmaceuticals.[36] Electron capture negative ionization (ECNI) excels for halogenated pollutants due to high electron affinity.[37] However, for PAC analysis, EI remains the most popular ionization technique.

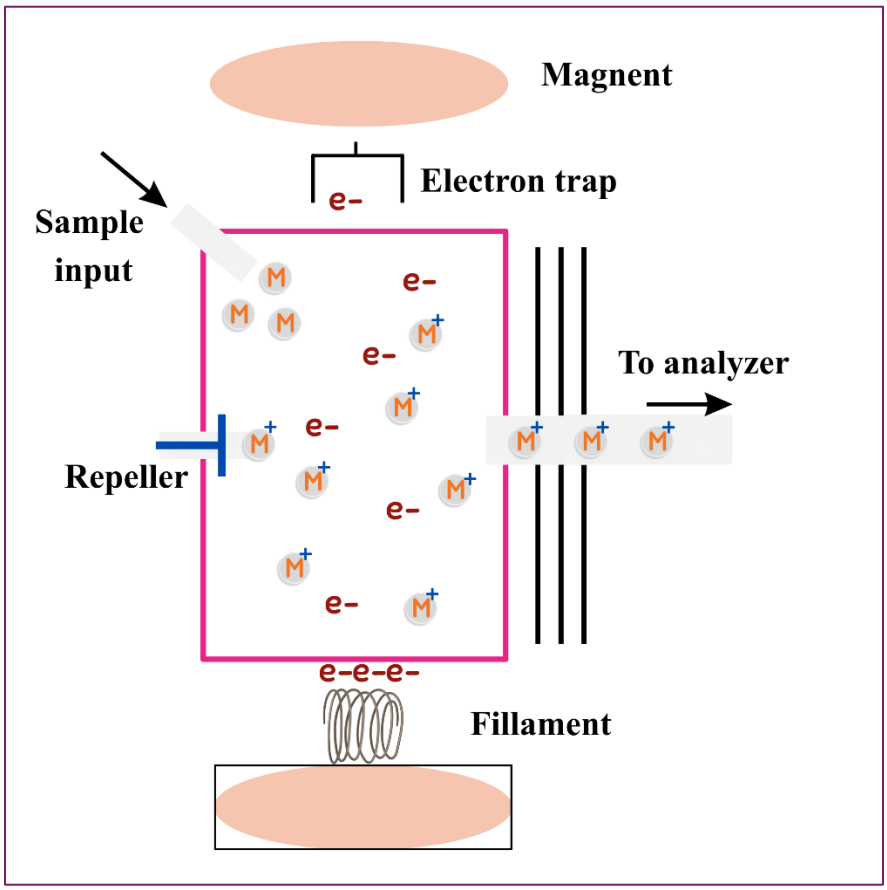



**Figure 2.1:** Schematic diagram of an electron ionization source used in mass spectrometry

Several types of mass analyzers have been developed, each offering distinct advantages in terms of resolution, mass accuracy, and scanning speed. Commonly used analyzers include quadrupole,[38] time-of-flight (TOF),[39] ion trap,[40] and Orbitrap systems.[33] Quadrupole analyzers are widely employed for routine quantitative analysis due to their robustness and selectivity. In contrast, TOF

and Orbitrap analyzers provide high-resolution and accurate mass measurements, enabling the determination of elemental compositions and the identification of unknown compounds.[41]

After separation, ions are detected and converted into electrical signals, which are processed to generate a mass spectrum. The resulting spectrum represents the relative abundance of ions as a function of their m/z values and serves as a unique fingerprint for compound identification. Advances in detector technology and data processing algorithms have significantly improved sensitivity, dynamic range, and spectral interpretation capabilities.

Coupling MS with chromatographic techniques was a major achievement in analytical chemistry, combining the high separation efficiency of chromatography with the structural identification capability of MS, making it one of the most powerful tools for analyzing complex mixtures.[36,42] Gas chromatography has been used in many environmental studies to identify organic pollutants.[43] In GC–MS, compounds are first separated in the chromatographic column and subsequently introduced into the mass spectrometer, where they are ionized, fragmented, and detected.

Due to its high sensitivity, selectivity, and ability to provide both qualitative and quantitative information, GC–MS has become a widely used technique in environmental analysis.[44] It is especially well-suited for the detection and characterization of volatile and semi-volatile organic compounds, including PACs, in complex environmental matrices such as sediments, water, and biota. Furthermore, the availability of extensive spectral libraries and high-resolution mass spectrometry enables both targeted and non-targeted analyses, facilitating the identification of known and emerging contaminants. [45]

### 2.2.3 Gas chromatography coupled with tandem mass spectrometry

Gas chromatography coupled with tandem mass spectrometry is a highly selective technique widely used in environmental analysis for trace-level detection of pollutants like PAHs, PCBs, dioxins, and PBDEs in complex matrices such as air particulates, sediments, and soils.[34,46,47] In GC-MS/MS with quadrupoles, mass filtering occurs in two stages, with ions passing through two quadrupoles, each consisting of four parallel hyperbolic or cylindrical rods forming a two-dimensional quadrupole field.[48] The theory of mass filtering through quadrupoles is as follows. Ions entering along the central axis experience a time-varying electric field from combined radio frequency (RF) and direct current (DC) voltages applied to opposite rod pairs. The resulting motion is described by the Mathieu equations:

$$\frac{d^2u}{d\xi^2} + (a_u + 2q_u\cos 2\xi)u = 0 \quad \frac{d^2v}{d\xi^2} + (a_v - 2q_v\cos 2\xi)v = 0$$

$$\text{where } \xi = \frac{\omega t}{2\sqrt{2}},\ a_u = a_v = \frac{8zeU}{mr_0^2\omega^2},\ \text{and } q_u = -q_v = \frac{4zeV}{mr_0^2\omega^2}$$

All the terms and symbols related to the equations are described below.

| | |
|---|---|
| $u, v$ | Radial coordinates in the quadrupole field (displacement in the x- and y-directions perpendicular to ion motion) |
| $\xi$ | Normalized time variable, defined as $\xi = \frac{\omega t}{2\sqrt{2}}$ |
| $a_u, a_v$ | Stability parameter (DC component), $a_u = a_v = \frac{8zeU}{mr_0^2\omega^2}$ |
| $q_u, q_v$ | Stability parameter (RF component), $q_u = -q_v = \frac{4zeV}{mr_0^2\omega^2}$ |
| $\omega$ | Angular frequency of the RF voltage ($\omega = 2\pi f$, where $f$ is RF frequency) |
| $t$ | Time |
| $U$ | DC voltage amplitude applied to quadrupole rods |

$V$ RF voltage amplitude applied to quadrupole rods.

$z$ Charge number of the ion (e.g., $z = 1$ for singly charged ions)

$e$ Elementary charge ($1.602 \times 10^{-19}$ C)

$m$ Mass of the ion

$r_0$ Field radius (characteristic dimension of quadrupole, distance from center to inner rod surface)

Stable ions trace bounded trajectories (defined by stability parameters. $a$ and $q$ within the stability diagram), reaching the detector; unstable ions are ejected.[48,49]

There are several operational modes in GC-MS/MS, such as full scan (FS), selected ion monitoring (SIM), product ion scan (PIS), and multiple reaction monitoring (MRM). In my research, I have used FS, PIS, and MRM modes. In FS mode, mass filtering does not occur at either quadrupole, and all ions produced by the ion source are directed to the detector, resulting in a scan of the sample.

In MRM, mass filtering occurs at two stages, based on defined m/z parameters, to select targeted compounds. Here, the first quadrupole (Q1) acts as a mass filter to select precursor ions/ parent ions from the GC-separated analytes. The central quadrupole (q2) serves as a collision cell for fragmentation via collision-induced dissociation (CID) with inert gas, and the third quadrupole (Q3) filters the product ions.

Product ion scan (PIS) is another operational mode I used in MS/MS, in which Q1 and Q2 are active while Q3 is not. This mode is generally used during method development to identify product ions. Here, the fragmentation patterns of different product ions at varying collision energies can be identified to select the best product ions for MRM analysis. A schematic diagram of the above-mentioned three operational modes of GC-MS/MS is shown in Figure 2.2.

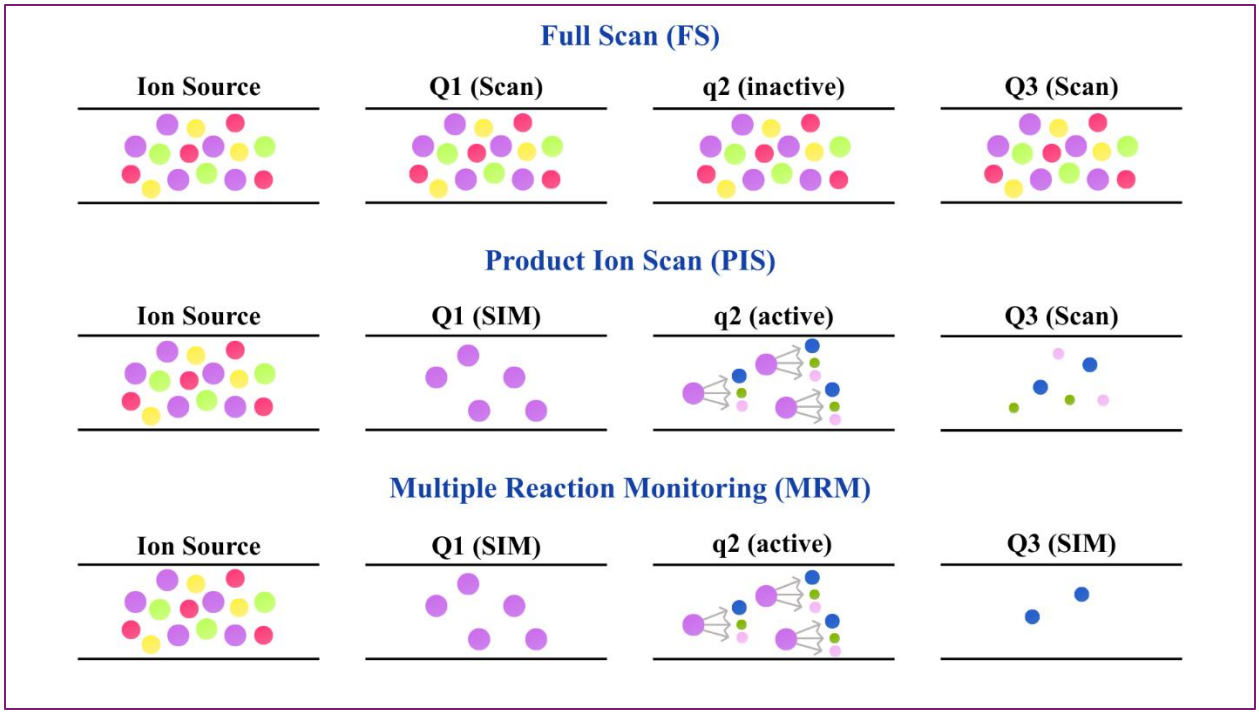



**Figure 2.2:** The three operational modes: full scan, product ion monitoring and multiple reaction monitoring used in GC-MS/MS in this study

In my work, an Agilent 7890 GC coupled to a triple-quadrupole Agilent 7000C mass spectrometer with an electron ionization (EI) source was used for GC-MS/MS analysis. I have developed two MRM methods for detecting 50 HPACs and Halo-HPACs. The method also include naphthalene-$d_8$, fluorene-$d_{10}$, phenanthrene-$d_{10}$, fluoranthene-$d_{10}$, pyrene-$d_{10}$, benz[a]anthracene-$d_{12}$, chrysene-$d_{12}$, benzo[b]fluoranthene-$d_{12}$, benzo[a]pyrene-$d_{12}$ and benzo[g,h,i]perylene-$d_{14}$ as the recovery internal standards (RIS) and anthracene-$d_{10}$ as the instrument performance internal standard (IPIS).

An Agilent J&W DB-5ms ultra-inert column (30 m × 0.25 mm, 0.25 µm) was used for chromatographic separation, with helium as the carrier gas at a constant flow rate of 1.2 mL/min. Samples (1 µL) were introduced in splitless mode using a PAL RSI 85 autosampler into an injector port maintained at 310 °C, equipped with a glass liner packed with wool. A splitless delay time of 0.75 minutes was applied. The oven temperature was held at 60 °C for 1 min then raised to 120 °C

at 35 °C min−1, further ramped up to 220 °C at 14 °C min−1, to 260 °C at 3 °C min−1 and held for 5 min, 300 °C at 10 °C min−1, and Finally to 310 °C at 50 °C $min^{-1}$. Both the transfer line and source temperatures were 300 °C.

The MRM methods used for PAH and APAH analysis were previously published by our group.[17] All MRM methods included a precursor ion and two product ions that acted as quantitation ions and confirmation ions. Table 2.1 summarizes all the analytes, their MRM transitions, and collision energies, in retention-time order.

**Table 2.1:** The targeted analytes and their multiple reaction monitoring parameters used in this study

| Targeted analytes | Acronym | PAC class | Internal standard | Parent Ion | Product Ion 1 | Product Ion 2 | Collision energy (eV) |
|---|---|---|---|---|---|---|---|
| Naphthalene-d8 | Nap-d8 | ISTD | - | 136 | 134 | 108 | 35 |
| Naphthalene | Nap | PAH | Nap-d8 | 127.9 | 102 | 77 | 45 |
| 2,3-Dibromothiophene | 2,3-Br2-BT | PASH | Nap-d8 | 242 | 163 | 161 | 15 |
| Quinoline | Qui | PANH | Nap-d8 | 129 | 102 | 76 | 25 |
| Isoquinoline | IQui | PANH | Nap-d8 | 129 | 102 | 76 | 35 |
| Indole | Ind | PANH | Nap-d8 | 117 | 90 | 89 | 35 |
| 2-Methylnaphthalene | 2-MeNap | APAH | Nap-d8 | 141.8 | 141 | - | 15 |
| 1-Methylnaphthalene | 1-MeNap | APAH | Nap-d8 | 141.8 | 141 | - | 15 |
| 4-Methylbenzothiophene | 4-MeBT | PASH | Nap-d8 | 147 | 146 | 121 | 25 |
| 2-Mequinoline | 2-MeQui | PANH | Nap-d8 | 143 | 128 | 115 | 25 |
| 3-Meisoquinoline | 3-MeIQui | PANH | Nap-d8 | 143 | 128 | 115 | 25 |
| 2,5-Dimethyl-1-phenylpyrrole | 2,5-DiMe-1-PP | PANH | Nap-d8 | 171 | 170 | 154 | 25 |
| 2-Chloroquinoline | 2-Cl-Qui | PANH | Flu-d10 | 163 | 128 | 101 | 25 |
| 5-Chlorobenzothiophene | 5-Cl-BT | PASH | Flu-d10 | 168 | 133 | 89 | 35 |
| 2,3-Dimethylbenzothiophene | 2,3-DiMeBT | PASH | Nap-d8 | 147 | 115 | 103 | 45 |
| C2-naphthalene | C2-Nap (1,4-DiMeNap) | APAH | Nap-d8 | 155.8 | 141 | - | 15 |
| Acenaphthylene d8 | Acy-d8 | ISTD | - | 160 | 158 | 156 | 25 |
| 2-Bromobenzothiophene | 2-Br-BT | PASH | Flu-d10 | 214 | 135 | 133 | 35 |
| Acenaphthylene | Acy | PAH | Acy-d8 | 151.9 | 151 | 150 | 35 |
| Acenaphthene d10 | Ace-d10 | ISTD | - | 162 | 160 | 158 | 15 |
| 2,3-Dimethylindole | 2,3-DiMeInd | PANH | Flt-d10 | 144 | 77 | 51 | 15 |
| Acenaphthene | Ace | PAH | Ace-d10 | 152.9 | 152 | 151 | 25 |
| 2,3-Dichlorobenzothiophene | 2,3-Cl2-BT | PASH | Nap-d8 | 202 | 167 | 123 | 25 |
| Dibenzofuran | DBF | PAOH | Flu-d10 | 168 | 139 | 129 | 25 |
| 2-Bromoquinoline | 2-Br-Qui | PANH | Flu-d10 | 128 | 101 | 77 | 25 |
| 5-Chloroindole | 5-Cl-Ind | PANH | Flu-d10 | 151 | 116 | 89 | 25 |
| C3-naphthalene | C3-Nap (2,3,5-TriMeNap) | APAH | Nap-d8 | 169.8 | 155 | 141 | 15 |
| Fluorene d10 | Flu-d10 | ISTD | - | 176 | 174 | 172 | 25 |
| Fluorene | Flu | PAH | Flu-d10 | 164.9 | 163 | 115 | 15 |
| 3-Bromoisoquinoline | 3-Br-IQui | PANH | Nap-d8 | 128 | 101 | 77 | 40 |
| 2,3,4-Trimethylbenzothiophene | 2,3,4-TriMeBT | PASH | Phe-d10 | 176 | 161 | 128 | 25 |
| 2-Methyldibenzofuran | 2-MeDBF | PAOH | Flu-d10 | 181 | 152 | 151 | 45 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 5-Bromoindole | 5-Br-Ind | PANH | Nap-d8 | 116 | 89 | 63 | 35 |
| C2-fluorene | C2-Flu (9-EtFlu) | APAH | Flu-d10 | 193.8 | 179 | | 35 |
| C4 naphthalene | C4-Nap (1,4,6,7-TetraMeNap) | APAH | Nap-d8 | 183.8 | 169 | 155 | 35 |
| C1-fluorene | C1-Flu (1-MeFlu) | APAH | Flu-d10 | 179.8 | 165 | | 15 |
| 2,3-Dibromobenzothiophene | 2,3-Br2-BT | PASH | Flu-d10 | 132 | 93 | 69 | 25 |
| Dibenzothiophene | DBT | PASH | Phe-d10 | 183.8 | 139 | | 15 |
| C3-fluorene | C3-Flu (9-PropFlu) | APAH | Flu-d10 | 208.1 | 179 | 165 | 15 |
| Phenanthrene d10 | Phe-d10 | ISTD | | 188 | 184 | 160 | 15 |
| Phenanthrene | Phe | PAH | Phe-d10 | 177.9 | 178 | 151 | 40 |
| Anthracene d10 | Ant-d10 | ISTD | | 188 | 184 | 160 | 40 |
| Anthracene | Ant | PAH | Ant-d10 | 177.9 | 178 | 151 | 15 |
| Benzo[h]quinoline | BhQui | PANH | Phe-d10 | 179 | 178 | 151 | 15 |
| 1,2,3,4-Tetrahydrocarbazole | 1,2,3,4-TetHCar | PANH | Phe-d10 | 143 | 128 | 115 | 35 |
| Acridine | Acr | PANH | Phe-d10 | 179 | 178 | 151 | 15 |
| 9-Methylcarbazole | 9-MeCar | PANH | Phe-d10 | 181 | 180 | 152 | 25 |
| Carbazole | Car | PANH | Phe-d10 | 167 | 166 | 138 | 15 |
| C4-fluorene | C4-Flu (9-ButFlu) | APAH | Flu-d10 | 222.1 | 193 | 179 | 25 |
| C1-dibenzothiophene | C1-DBT (2-MeDBT) | PASH | Phe-d10 | 197.8 | 197 | - | 15 |
| 3-Methylbenzo[f]quinoline | 3-MeBfQui | PANH | Phe-d10 | 193 | 165 | 151 | 45 |
| 3-Methylphenanthrene | 3-Phe | APAH | Phe-d10 | 191.8 | 191 | - | 25 |
| 2-Methylphenanthrene | 2-Phe | APAH | Phe-d10 | 191.8 | 191 | - | 40 |
| 9-Methylphenanthrene | 9-Phe | APAH | Phe-d10 | 191.8 | 191 | - | 40 |
| 1-Methylphenanthrene | 1-Phe | APAH | Phe-d10 | 191.8 | 191 | - | 40 |
| 2-Methylacridine | 2-MeAcr | PANH | Phe-d10 | 193 | 192 | 191 | 40 |
| 9-Chloroacridine | 9-Cl-Acr | PANH | Phe-d10 | 213 | 178 | 151 | 25 |
| C3-dibenzothiophene | C3-DBT (4-PropDBT) | PASH | Phe-d10 | 225.8 | 211 | 197 | 25 |
| 1,3-Dimethylphenanthrene | 1,3-DiMePhe | APAH | Phe-d10 | 205.8 | 191 | - | 25 |
| 3,6-Dimethylphenanthrene | 3,6-DiMePhe | APAH | Phe-d10 | 205.8 | 191 | - | 25 |
| 2,6-Dimethylphenanthrene | 2,6-DiMePhe | APAH | Phe-d10 | 205.8 | 191 | - | 25 |
| 1,7-Dimethylphenanthrene | 1,7-DiMePhe | APAH | Phe-d10 | 205.8 | 191 | - | 15 |
| 1,8-Dimethylphenanthrene | 1,8-diMePhe | APAH | Phe-d10 | 205.8 | 191 | - | 15 |
| 3,6-Dimethylcarbazole | 3,6-DiMeCar | PANH | Flt-d10 | 180 | 179 | 152 | 35 |
| C2-dibenzothiophene | C2-DBT (1,2-DiMeDBT) | PASH | Phe-d10 | 211.8 | 197 | - | 15 |
| 2-Bromodibenzothiophene | 2-Br-DBT | PASH | Phe-d10 | 264 | 185 | 183 | 15 |
| 9-Bromoacridine | 9-Br-Acr | PANH | Phe-d10 | 177.9 | 151 | 125 | 15 |
| Fluoranthene d10 | Flt-d10 | ISTD | | 212 | 210 | 208 | 15 |
| Fluoranthene | Flt | PAH | Flt-d10 | 201.9 | 201 | 200 | 25 |
| Pyrene d10 | Pyr-d10 | ISTD | | 212 | 210 | 208 | 15 |
| Pyrene | Pyr | PAH | Pyr-d10 | 201.9 | 201 | 200 | 15 |
| C4-dibenzothiophene | C4-DBT (4,6-DiEtDBT) | PASH | Flt-d10 | 239.9 | 224.5 | 211 | 15 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Benzo[b]naphtho[2,3-d]furan | BbNF(2,3-d) | PAOH | Phe-d10 | 188.8 | 187 | 163 | 15 |
| 3-Tert-butyl-9H-carbazole | 3-tBut-9H-Car | PANH | Pyr-d10 | 208 | 180 | 167 | 15 |
| C3-phenanthrene | C3-Phe (1,2,6-TriMePhe) | APAH | Pyr-d10 | 220 | 205 | 191 | 15 |
| Retene | Ret | APAH | Phe-d10 | 219 | 204 | - | 15 |
| C1-pyrene | C1-Pyr (1-MePyr) | APAH | Phe-d10 | 215.8 | 216 | 215 | 25 |
| 9-Phenylcarbazole | 9-Ph-Car | PANH | Pyr-d10 | 243 | 241 | 166 | 15 |
| C4-phenanthrene | C4-Phe (1,2,6,9-TetMePhe) | APAH | Pyr-d10 | 233.8 | 219 | 205 | 25 |
| C2-pyrene | C2-Pyr (1-EtPyr) | APAH | Phe-d10 | 229.8 | 215 | - | 15 |
| 3,6-Dichloro-9h-carbazole | 3,6-Cl2-Car | PANH | Pyr-d10 | 234.8 | 200 | 164.3 | 45 |
| 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | 7,8,9,10-H4-BbNT(2,3-d) | PASH | Pyr-d10 | 210 | 208 | 184 | 45 |
| 2,8-Dibromodibenzothiophene | 2,8-Br2-DBT | PASH | Pyr-d10 | 182 | 138 | 138 | 45 |
| Benz[c]acridine | BcAcr | PANH | Pyr-d10 | 229 | 228 | 227 | 45 |
| Benzo [b]naphtho[1,2-d]thiophene | BbNT(1,2-d) | PASH | Chr-d12 | 234.1 | 201 | 189 | 15 |
| Benzo [b]naphtho[2,3-d]thiophene | BbNT(2,3-d) | PASH | Pyr-d10 | 234.1 | 201 | 189 | 15 |
| C3-pyrene | C3-Pyr (1-PropPyr) | APAH | Pyr-d10 | 243.8 | 229 | 215 | 35 |
| Benz(a)anthracene d12 | BaA-d12 | ISTD | - | 240 | 236 | 212 | 15 |
| Benz(a)anthracene | BaA | PAH | BaA-d12 | 227.9 | 226 | 202 | 15 |
| Chrysene d12 | Chr-d12 | ISTD | - | 240 | 236 | 212 | 15 |
| Chrysene | Chr | PAH | Chr-d12 | 227.9 | 226 | 202 | 15 |
| 11H-Benzo[a]carbazole | 11H-BaCar | PANH | Chr-d12 | 217 | 216 | 188.8 | 35 |
| 7H-benzo[c]carbazole | 7H-BcCar | PANH | Chr-d12 | 217 | 216 | 188.8 | 15 |
| 6-Methylbenzo [b]naphtho[2,1-d]thiophene | 6-MeBbNT(2,1-d) | PASH | Chr-d12 | 248 | 247 | 245 | 15 |
| C4-pyrene | C4-Pyr (1-ButPyr) | APAH | BaP-d12 | 257.9 | 243 | 228 | 25 |
| 3,6-Dibromocarbazole | 3,6-Br2-Car | PANH | Pyr-d10 | 324.8 | 246 | 243.9 | 25 |
| C1-chrysene | C1-Chr (5-MeChr) | APAH | BaP-d12 | 241.8 | 239 | - | 15 |
| 2-Phenyldibenzothiophene | 2-Ph-DBT | PASH | Chr-d12 | 259.9 | 257 | 256 | 25 |
| C2-chrysene | C2-Chr (5-EtChr) | APAH | BghiP-d12 | 255.8 | 241 | - | 35 |
| C3-chrysene | C3-Chr (6-PropChr) | APAH | Chr-d12 | 269.8 | 241 | - | 35 |
| Benzo(b)fluoranthracene d12 | BbF-d12 | ISTD | - | 264 | 260 | 236 | 15 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Benzo(b)fluoranthracene | BbF | PAH | BbF-d12 | 251.9 | 250 | 226 | 40 |
| Benzo(k)fluoranthracene d12 | BkF-12 | ISTD | - | 264 | 260 | 236 | 40 |
| Benzo(k)fluoranthracene | BkF | PAH | BbF-d12 | 251.9 | 250 | 226 | 40 |
| Dinaphtho[1,2-b;1',2'-d]furan | DNF[1,2;1,2-d] | PAOH | BkF-d12 | 267.8 | 239 | 237 | 40 |
| C4-chrysene | C4-Chr (6-ButChr) | APAH | BghiP-d12 | 283.8 | 269 | 255 | 30 |
| Benzo(a)pyrene d12 | BaP-d12 | ISTD | - | 264 | 260 | 236 | 30 |
| Benzo(a)pyrene | BaP | PAH | BaP-d12 | 251.9 | 250 | 226 | 25 |
| C1-benzo(a)pyrene | C1-BaP (6-MeBaP) | APAH | BaP-d12 | 266 | 263 | - | 15 |
| Dibenz[c,h]acridine | DBchAcr | PANH | BaP-d12 | 279 | 278 | 277 | 15 |
| 13(H)-Dibenzo[a,i]carbazole | 13(H)-DBaiCar | PANH | BghiP-d12 | 267 | 266 | 239 | 20 |
| C2-benzo(a)pyrene | C2-BaP (7,10-DiMeBaP) | APAH | BghiP-d12 | 280 | 265 | - | 40 |
| Indeno(1,2,3-cd)pyrene d12 | IcdP-d12 | ISTD | - | 288 | 288 | 275 | 15 |
| Indeno(1,2,3-cd)pyrene | IcdP | PAH | IcdP-d12 | 275.8 | 275.8 | - | 15 |
| Dibenz(a,h)anthracene d14 | DBahA-d14 | ISTD | - | 292 | 292 | 277.9 | 45 |
| Dibenz(a,h)anthracene | DBahA | PAH | DBahA-d14 | 278 | 278 | - | 40 |
| Benzo(g,h,i)perylene d12 | BghiP-d12 | ISTD | - | 288 | 288 | 277.9 | 45 |
| Benzo(g,h,i)perylene | BghiP | PAH | BghiP-d12 | 275.9 | 276 | - | 40 |

### 2.2.4 Two-dimensional gas chromatography coupled with high resolution time of flight mass spectrometry

The complex nature of environmental samples often results in poor compound separation when a one-dimensional column is used.[50] The compound classes, like PACs, have numerous derivatives and isomers based on alkylation, halogenation, and heteroatom position. When APAHs are considered, the number of possible theoretical isomers substantially increases with the degree of alkylation. For example, for phenanthrene from C1 to C4, the number of theoretically possible isomers increases from five to 110.[51] Since isomers can not be mass resolved, the maximum possible separation needs to be achieved using GC.

Comprehensive two-dimensional GC (GCxGC) is a technique that uses two orthogonal columns to analyze complex mixtures, providing advanced separation and increased resolution.[52] Typical GCxGC instruments consist of three major sections. Primary non-polar column that is in the primary oven, modulator, and secondary polar column that runs through the modulator and the secondary oven.[53] A schematic diagram of the GCxGC instrument is shown in Figure 2.3.

A sample is injected into the primary column (usually a 30m column), where separation occurs based on analyte volatility. Then, the separated peaks are directed to the modulator, which continuously captures small fractions of the 1D peak (~25%), refocuses them, and rapidly injects them into the secondary column at a frequency of 0.1–1.0 Hz.[54] The secondary column is usually about 2m long and coated with a polar stationary phase. Hence, the compounds are further separated by polarity. [55]

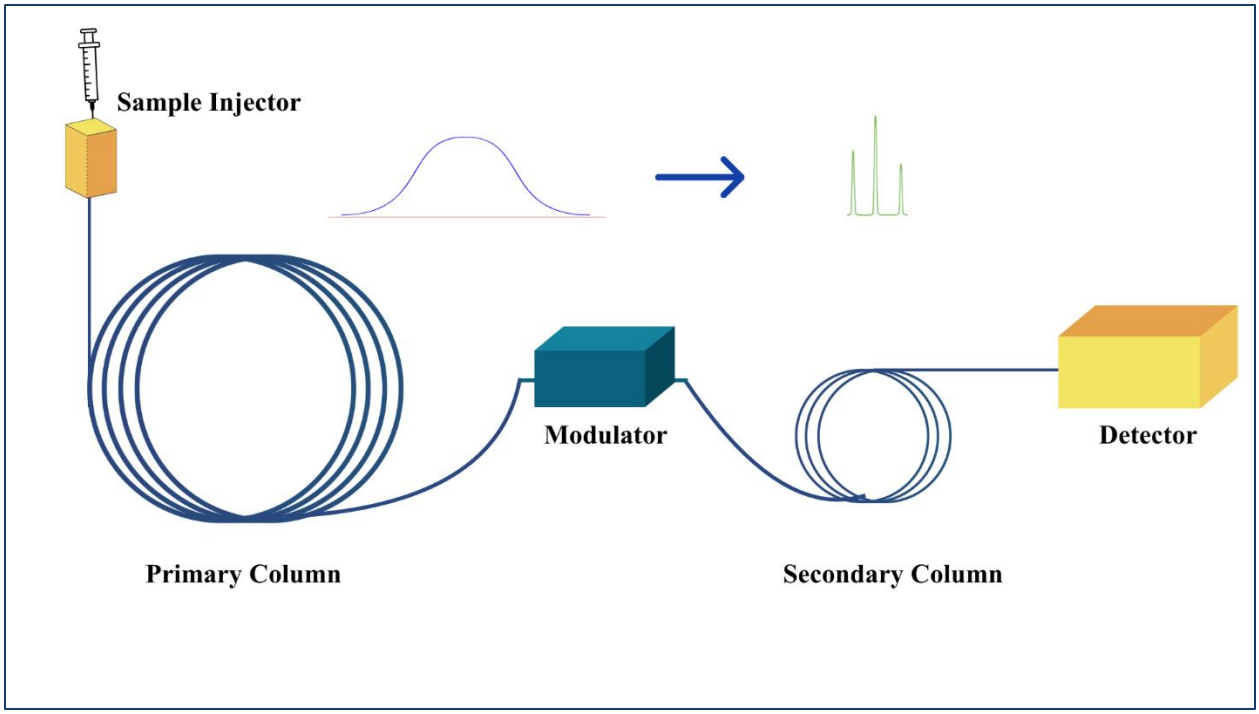



**Figure 2.3:** Schematic diagram of a two-dimensional gas chromatography setup

Low-resolution MS systems typically provide unit mass resolution and are therefore unable to distinguish between isobaric compounds.[56] Although MS/MS instruments offer higher selectivity than conventional low-resolution systems, they are generally limited to identifying targeted compounds for which authentic standards are available. However, environmental samples contain PACs as complex mixtures, and reliance solely on targeted analysis may not provide a comprehensive understanding of their overall environmental impact.[57]

High-resolution mass spectrometry (HRMS) is an advanced analytical technique that enables measurement of ion masses with very high accuracy, typically at the parts-per-million (ppm) level or lower.[58] Unlike low-resolution mass spectrometry, HRMS can resolve compounds with very small mass differences, increasing mass resolution. This enables the determination of exact molecular formulas and increases confidence in compound identification.

Common types of high-resolution mass analyzers include time-of-flight (TOF), Orbitrap, and Fourier transform ion cyclotron resonance (FT-ICR) instruments.[56] Among these, time-of-flight analyzers are widely used in comprehensive GC×GC applications due to their high acquisition rates and ability to handle the large volumes of data generated by complex separations.

Overall, HRMS plays a critical role in modern analytical chemistry by enhancing selectivity, sensitivity, and identification capability, particularly in studies involving trace-level contaminants and complex mixtures such as PACs in environmental matrices.

Time-of-flight mass spectrometry measures the m/z of ions by recording the time they take to travel a fixed distance through a field-free drift tube after acceleration by an electric field. [56] These instruments mainly consist of three components. First, an ion source that generates ions in short pulses. Then there is an acceleration region that imparts kinetic energy to the ions by providing a voltage. Then there is a field-free drift tube, called a flight tube, in which ions separate based on their velocity before detection.[59,60] Figure 2.4 shows a schematic diagram of a TOF-MS.

In TOF-MS, ions are generated in a pulsed ion source and accelerated to the same kinetic energy. In the flight tube, lighter ions travel faster and arrive at the detector first. The flight time (t) is proportional to the m/z values of ions, and based on t, the masses get resolved.[60]

Key advantages of TOF include unlimited $m/z$ range, high speed (>$10^3$ Hz acquisition), and full-spectrum sensitivity. The resolution of these mass spectrometers can be improved by increasing the path length or by using a reflectron. This compensates for the initial spread of kinetic energy by retarding or reflecting faster ions deeper into the material.[61,62]

The equations related to the theory of operation of TOF-MS are described below.

$$E_k = \frac{1}{2}mv^2 = zeU$$

$$v = \sqrt{\frac{2ezU}{m}}$$

$$t = \frac{L}{v} = L\sqrt{\frac{m}{2ezU}}$$

$$R = \frac{m}{\Delta m} = \frac{1}{2}\frac{t}{\Delta t}$$

| | |
|---|---|
| $v$ | Ion velocity |
| $E_k$ | Kinetic energy |
| $t$ | Flight time (TOF) |
| $L$ | Drift tube length |
| $m$ | Ion mass |
| $z$ | Ion charge number |
| $e$ | Elementary charge |
| $U$ | Acceleration voltage |
| $R$ | Resolving power (FWHM) |
| $\Delta t$ | Time spread |

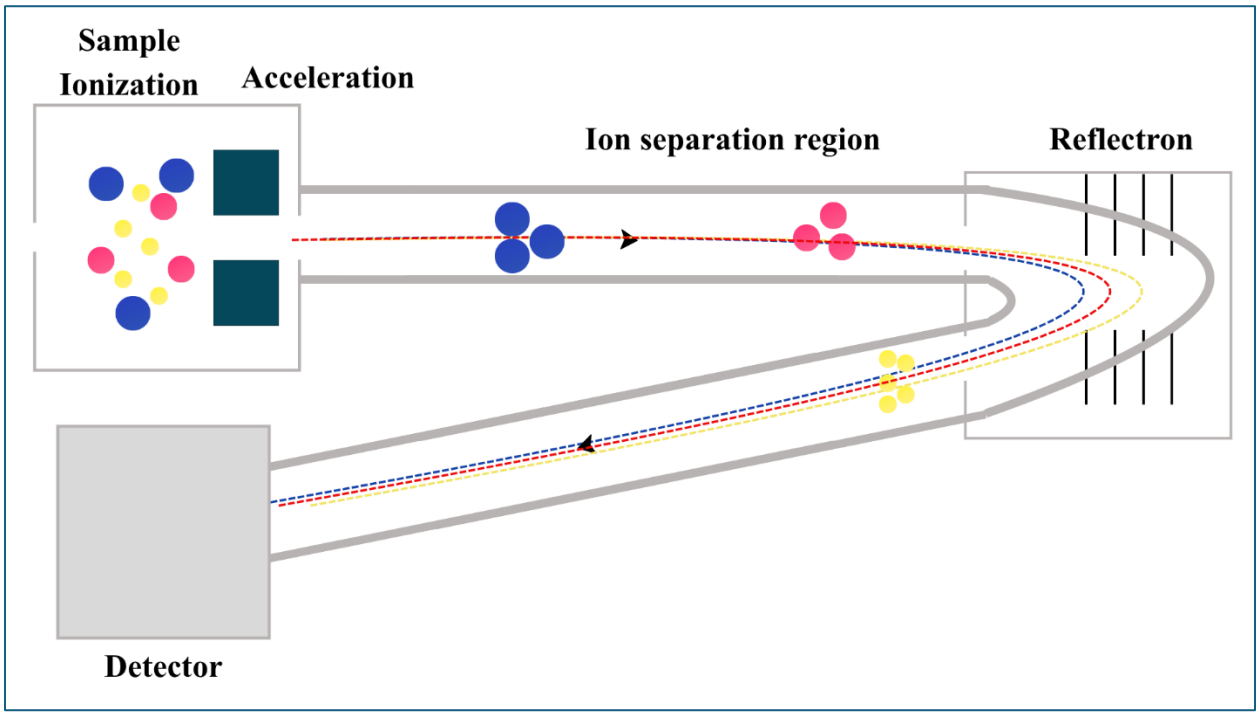



**Figure 2.4:** Schematic diagram of a high-resolution time-of-flight mass spectrometer

In my study, non-targeted analysis was carried out on Lake Ontario sediment and Lake Erie fish samples to identify PACs using GCxGC-HRTOF-MS. High-resolution mass spectrometric analysis was performed using an Agilent 7890A gas chromatograph coupled with a Pegasus GC-HRT 4D (LECO, St. Joseph, MI, USA) high-resolution time-of-flight mass spectrometer, following a previously developed and validated method.[55] Injections of 2 µL were carried out in splitless mode at an injector temperature of 300 °C.

Chromatographic separation was achieved using a two-dimensional GC configuration, with an Rxi-5Sil MS column (60 m × 0.25 mm, 0.25 µm; Restek, Bellefonte, PA, USA) as the primary column and an Rxi-17Sil MS column (2 m × 0.25 mm × 0.25 µm; Restek) as the secondary column, connected via an SGE micro-union. Helium was used as the carrier gas.

The primary oven temperature was programmed from 55 °C (held for 2 minutes), increased to 110 °C at 10 °C/min, then to 210 °C at 3 °C/min, and finally to 300 °C at 8 °C/min, with a final hold of 13 minutes. The secondary oven and thermal modulator temperatures were maintained at 5 °C and 15 °C above the primary oven temperature, respectively.

The mass spectrometer operated at a resolving power was set as full width at half maximum (FWHM) of 25,000, determined based on the peak width of m/z 218.9851 of perfluorotributylamine (PFTBA). During analysis, PFTBA was continuously introduced into the ion source to serve as an internal standard for real-time mass calibration. Solvent blanks were analyzed after every five sample injections to monitor potential carryover and ensure data quality.

**2.3 DATA PROCESSING AND INTERPRETATION**

Another major step in any analytical procedure is the accurate processing and interpretation of data. This step turns raw analytical signals into meaningful results, and that is essential for

informed decision-making. This step includes peak detection, deconvolution, compound identification, and quantification, often supported by spectral libraries and in-house databases. The application of appropriate statistical and chemometric tools further enables the identification of patterns, trends, and relationships within the data, enhancing the interpretability and confidence of the findings.

In the targeted analysis, spectral data were processed using Agilent MassHunter Quantitative analysis software (2015, version B.07.01). Using the spectral area data for the target analytes and internal standards, final concentrations were calculated by the internal standard method.[63,64] The concentration data were analyzed using a series of statistical and chemometric tools to identify the source, fate, and distribution of PACs in environmental systems.

### 2.3.1 Statistical analysis

Analytical chemistry is a science that relies heavily on quantitative measurements, and analytical data are inherently subject to variability arising from instrumental limitations, sample heterogeneity, and experimental conditions. Statistics can be used to evaluate these variabilities. Therefore, statistics provides the necessary tools to design analytical methods, evaluate data quality, and draw conclusions from acquired raw data.[65]

Beyond basic statistical techniques, modern analytical data generated by advanced instrumentation often require the integration of multiple disciplines, such as mathematics, statistics, machine learning, artificial intelligence, and electronics, for accurate interpretation. The application of these combined approaches to the analysis and interpretation of chemical data is referred to as chemometrics (see Figure 2.4).[66]

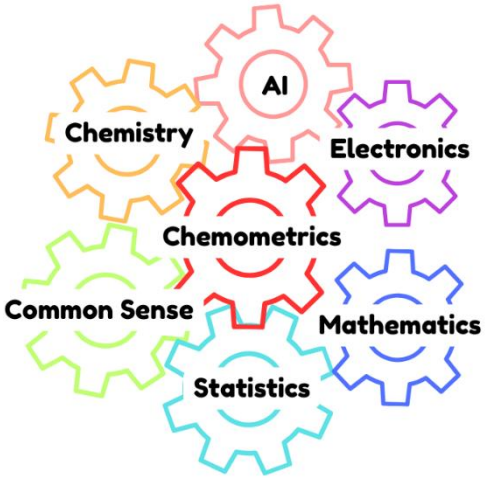



**Figure 2.5:** A diagram that summarizes how chemometrics are interlinked with different subject areas

Over the course of my thesis, I used several basic statistics and chemometrics tools in steps such as establishing calibration ranges for HPACs, evaluating method performance characteristics, and interpreting thousands of concentration data points for PACs. In this section, those statistical tools and their application are described.

#### 2.3.1.1 Regression analysis and calibration models

Regression analysis is one of the most frequently used statistical techniques in analytical chemistry. Regression analysis models and quantifies the relationship between a dependent variable (response, y variable) and one or more independent variables (predictors, x variables). There are several types of regression analysis methods, and in analytical chemistry, linear regression, nonlinear regression, weighted regression, and multiple linear regression are among the most used.[67,68]

For years, linear regression has been the most widely used technique in analytical chemistry.[69] The simplest form of linear regression is the ordinary least squares regression (OLSR): $y = c + \mathrm{m}x$,

where c is the intercept and m is the slope. This model is constructed by minimizing squared residuals between observed and predicted values.[70,71] However, to fit a linear model to a data set, the data should be normally distributed, have constant variance (homoscedasticity), and be independent of errors, enabling quantification of unknowns by solving for the dependent variables.[67] The goodness of fit is usually assessed via $R^2$, and the statistical significance of the model is assessed using the p-value ($p < 0.05$).[72,73] The importance of reporting the p-value is highlighted in the tropic magnification factor (TMF) analysis in Chapter 6.

In any quantitative analytical method, calibration curves are essential to produce accurate measurements.[74] Linear regression is the most common model used in constructing the calibration curve. However, it is important to evaluate the model capabilities beyond the $R^2$ values. Even though the $R^2$ value is close to 1, most analytical data can exhibit heteroscedasticity, in which the variance increases with concentration.[72] This can be evaluated by constructing a residual plot, and Chapter 3 describes how the HPACs' linear calibration data were found to be heteroscedastic.

For this reason, in my thesis, the average response factor (ARF, see equation below) model was used to calibrate HPACs.[7] The residual plots for the two models showed that, for most targeted analytes, the ARF model is more suitable for calibration data than the linear model, as large positive or negative biases were observed.

$$Average\ r_f = \sum \left(\frac{electronically\ integrated\ area\ of\ analyte}{\frac{electronically\ integrated\ area\ of\ d-PAH}{Concentration\ of\ analyte\ in\ standard}}\right)/n$$

where $n$ = the number of standards injected

#### 2.3.1.2 Normal distributions and environmental sample analysis

The normal distribution, also known as the Gaussian distribution, is one of the most important statistical concepts in analytical chemistry. In a normal distribution, data are distributed symmetrically around a central mean, forming a characteristic bell-shaped curve. A normal distribution can be fully described by two parameters: the mean (μ) and the standard deviation (σ). The mean is the central value of the distribution, and the standard deviation describes the spread or variability of the data around the mean. In a perfectly normal distribution, the mean, median, and mode are equal, and the data are symmetrically distributed about the mean.[75]

One of the key features of the normal distribution is its predictable spread. Approximately 68% of values lie within one standard deviation (±1σ) of the mean, while 95% within two standard deviations (±2σ), and 99.7% within three standard deviations (±3σ). This property is widely used in analytical chemistry to assess data quality and identify outliers.[76]

Many statistical tests used in analytical chemistry, such as correlation, factor analysis, and t-tests, assume that the data are normally distributed.[77] However, environmental data generally does not follow a normal distribution. Those data exhibit right-skewed distributions with heavy tails, outliers, and two or more peaks (multimodality) due to multiple underlying causes, such as varying sources, spatial heterogeneity, detection limits, and measurement imprecision.[75,78] The analysis of non-normal data using a statistical method that assumes normality biases means, variances, confidence intervals, and p-values, leading to invalid inferences.[79] When analytical data deviate significantly from normality, alternative statistical approaches or data transformations are applied.[80]

Therefore, in environmental sample analysis, it is important to check whether the concentration data follow a normal distribution to identify appropriate statistical methods. The Shapiro-Wilk test, the Kolmogorov-Smirnov test, the Anderson-Darling test, and the Chi-squared ($\chi^2$) test are some of the formal tests used to test the normality of data.[76]

In my thesis, to check the normality of the concentration data, the Shapiro-Wilk test was used. This test is particularly suitable for small to moderate sample sizes and is commonly applied in analytical chemistry before performing parametric statistical analyses. The test evaluates the null hypothesis that the data are normally distributed. It does so by comparing the order and spacing of the sample values to those expected from a theoretical normal distribution.[81] The result of the test is a test statistic along with a corresponding p-value, which indicates the likelihood of the data deviating from normality. If the p-value is greater than a chosen significance level (typically 0.05), the null hypothesis is not rejected, suggesting that the data do not significantly differ from a normal distribution. Conversely, a p-value less than 0.05 indicates that the data deviates significantly from normality. [82]

Before assessing non-normal data with parametric analyses that assume normality, data transformation methods are applied. There are several methods available for data transformation, and in this study, I have used lognormal, Hellinger, or percent normal transformations.

In a lognormal transformation, the log10 of the data is calculated to reduce right-skewness. It compresses large values more than small ones, making highly skewed environmental or concentration data more normally distributed and stabilizing variance.[83] In the percent-normal transformation, the values are divided by the sample total (relative abundance). The Hellinger transformation is commonly used in multivariate analysis of compositional data. It involves taking

the square root of relative abundances. This reduces the dominance of highly abundant variables and makes the concentration data suitable for methods such as PCA that rely on Euclidean distances.[84,85]

#### 2.3.1.3 Analysis of variance and method performance evaluation

Analysis of Variance (ANOVA) is a widely used statistical test that evaluates variance among three or more groups and determines whether there is a significant difference among the group means. The method tests the null hypothesis that all group means are equal. A significant result (typically $p < 0.05$) indicates that at least one group mean differs from the others, but it does not specify which groups differ.[86,87] Post hoc tests (e.g., Tukey's test) are often applied to identify specific group differences.[88,89]

ANOVA assumes that the data are normally distributed, homoscedastic, and independent. However, in most environmental samples, the concentration data is neither normally distributed nor homoscedastic. Therefore, in such cases, Welch's ANOVA is applied to the transformed data.[90]

Welch's ANOVA is a modified version of standard ANOVA that is used when the assumption of equal variances is violated. Unlike classical ANOVA, Welch's ANOVA does not assume homogeneity of variances and is therefore more robust when group variances and sample sizes are unequal. Like standard ANOVA, Welch's ANOVA tests whether there are significant differences between group means. However, it uses adjusted degrees of freedom and a weighted approach to account for unequal variances. As a result, it provides more accurate p-values and better controls the Type I error (false positives) rate when the assumption of homogeneity of variances is violated.[90,91]

In my thesis, standard ANOVA and Welch's ANOVA, along with post hoc tests, were applied in several instances. One such situation involved evaluating method performance characteristics, such as intermediate precision and ruggedness. In the intermediate precision evaluation step, a one-way ANOVA test was used to determine whether there was a significant difference in analyte recoveries across extractions and purifications performed in different batches on different days. In this case, ANOVA was the recommended evaluation tool per standard guidelines, such as the Eurachem Guide - The Fitness for the Purpose of Analytical Methods.[92]

#### 2.3.1.4 Correlation analysis

Correlation analysis evaluates the strength and direction of the relationship between two variables. In my study, two correlation analyses were conducted to determine the sources of PACs, as described in Chapter 4. Those correlation analyses are the Pearson Correlation Coefficient and Spearman's Rank Correlation analyses.

The Pearson Correlation Coefficient measures the strength of a linear relationship between two continuous variables. It assumes that the data are normally distributed and that the relationship is linear. The coefficient (r) ranges from -1 to +1, where values close to ±1 indicate strong correlations.[93,94] The Spearman's Rank Coefficient evaluates the strength of a monotonic relationship based on ranked data rather than raw values. It does not require normal distribution and is less sensitive to outliers.[95,96]

#### 2.3.1.5 Multivariate analysis and the environmental distribution of polycyclic aromatic compounds

The analysis of complex classes of environmental pollutants, such as PACs, generates large and multidimensional datasets. Interpreting such data using basic statistical measures like mean, standard deviation, and median is often insufficient to capture the full extent of their environmental significance.[97] Therefore, more advanced approaches to high-dimensional data analysis are required.[65]

Multivariate analysis is a branch of statistics that focuses on analyzing high-dimensional data, enabling the exploration of relationships and interactions among multiple variables simultaneously.[98] A variety of multivariate techniques are available, and in this study, four methods were employed: principal component analysis (PCA), uniform manifold approximation and projection (UMAP), redundancy analysis (RDA), and hierarchical cluster analysis (HCA). These techniques are introduced (Table 2.2) below and further explored in later chapters.

Principal component analysis is an unsupervised dimensionality reduction technique that transforms a large set of correlated variables into a smaller set of uncorrelated variables, called principal components.[99,100] These components capture the maximum variance in the data, allowing visualization of patterns, trends, and groupings.

Uniform manifold approximation and projection is an unconstrained, non-linear dimensionality reduction method that preserves both local and global data structure.[101] It is particularly effective for visualizing complex, high-dimensional datasets by projecting them into two or three dimensions. Compared to PCA, UMAP can better capture non-linear relationships, making it useful for identifying clusters and subtle patterns in environmental data.[102] In my study, I used

PCA and UMAP for source fingerprinting of PACs in Lake Erie sediments and to identify basin- and species-dependent variations in PACs in Lake Erie fish.

Redundancy analysis is a constrained multivariate technique that combines regression and ordination. It is used to assess how much of the variation in a set of response variables is explained by a set of explanatory variables, also called predictors.[103,104] In my study, RDA was used to assess how PAC concentrations in different species vary with other factors, such as stable isotope values, weight, and length of fish samples.

Hierarchical cluster analysis (HCA) is a statistical technique used to group samples or variables based on their similarity. In HCA, a distance or dissimilarity matrix is first calculated to quantify the differences between samples. A linkage method is then applied to iteratively merge similar samples or clusters. This process results in a dendrogram in which more similar samples are grouped at shorter distances, allowing for intuitive visualization of relationships within the data.[105,106] In this study, I have used Bray–Curtis distance and the average-linkage method to perform HCA to identify basin-dependent clusters in Lake Erie fish.

**Table 2.2:** Characteristics of the four multivariate approaches used in this study

| Aspect | Principal Component Analysis (PCA) | Uniform Manifold Approximation and Projection (UMAP) | Redundancy Analysis (RDA) | Hierarchical Cluster Analysis (HCA) |
|---|---|---|---|---|
| **Type** | Unsupervised | Unsupervised | Supervised (constrained) | Unsupervised |
| **Main purpose** | Reduce dimensionality, identify variance patterns | Reduce dimensionality, Visualize complex, non-linear structures | Reduce dimensionality, explain variation using predictors | Group similar samples. |
| **Relationship type** | Linear | Non-linear | Linear (constrained by predictors) | Distance based |
| **Input data** | High-dimensional numerical data | High-dimensional numerical data | High-dimensional response variables and predictors | High-dimensional numerical data |
| **Output** | Principal components | Low-dimensional variables | Low-dimensional variables driven by predictors | Dendrogram explaining clusters |
| **Interpretability** | High (loadings show variable influence) | Moderate (less direct interpretation) | High (links variables to predictors) | Moderate (clusters but no direct variable contribution) |
| **Handles complex data** | Moderate | High | Moderate | Moderate |
| **Sensitivity to scaling** | High (requires standardization) | Moderate | High | High |
| **The use in this study** | Identify trends and dominant variables (PACs) | Detect clusters and subtle patterns driven by PACs | Link PAC distributions to environmental factors | Classify samples based on PAC concentrations |

### 2.3.2 Non-targeted data analysis for the identification of polycyclic aromatic compounds

Non-targeted analysis is an analytical approach that detects and identifies a broad range of chemicals in complex environmental samples without prior selection of specific target compounds.[107] Non-targeted analysis workflows typically involve extraction, chromatography, and HRMS for exact-mass determination, followed by data processing for feature detection, alignment, and tentative identification.[108,109]

Since NTA does not use authentic standard comparisons for identification, as in targeted analysis, it relies on other identification tools such as library data, mass defect calculations, and relative retention times. The common challenges of NTA include matrix interferences, data overload, and quantification uncertainty.[110,111] However, in PAC, I faced another major challenge. Besides PAHs, most PAC classes are understudied, and therefore, library data for most possible PACs were not available in reference libraries. Therefore, in my research, I created a workflow to exclusively identify PACs by combining available standard data with PAC chemical characteristics using the R program. The detailed description of the workflow is provided in Chapter 8.

## 2.4: REFERENCES


(1) Idowu, I. G.; Dereviankin, M.; Halldorson, T.; Vitharana, N.; Harbicht, L.; Helm, P.; Kleywegt, S.; Marvin, C.; Sandau, C. D.; Tomy, G. T. Source Characterization of Polycyclic Aromatic Hydrocarbons in Urban Sediments: Insights from the Contribution of Coal Tar, Sealants, and Other Related Industrial Materials in Lake Ontario. *Science of the Total Environment* **2025**, 1009. https://doi.org/10.1016/j.scitotenv.2025.180967.

(2) Zhang, X.; Bowman, D. T.; Diamond, M. L.; Helm, P.; Jobst, K. J.; Hao, C.; Kleywegt, S.; Zhang, Z. F.; Marvin, C.; Zhang, X. Contribution of Coal Tar Sources to Polycyclic Aromatic Compounds and Associated Ecological Risk in Lake Ontario Sediments: Inference from a Novel Marker. *Environ. Sci. Technol*. **2025**. https://doi.org/10.1021/acs.est.4c10360.

(3) Kellems, B.; Fabian, K.; Eng, P.; Orchard, B.; Parmelee, R.; Bonkoski, B.; Spadaro, P.; Beaver, J.; Fitzgerald, W. P.; Santiago, R.; Rozenberg, W. Randle Reef Sediment Remediation Project; **2013**.

(4) Heuvel, C. E.; Zhao, Y.; Fisk, A. T. Influence of Spatial and Temporal Variation on Establishing Stable Isotope Baselines of Δ15N, Δ13C, and Δ34S in a Large Freshwater Lake. Freshw. *Biol.* **2023**, 68 (5), 806–820. https://doi.org/10.1111/fwb.14065.

(5) Heuvel, C. E.; Zhao, Y. M.; Fisk, A. T. Food Web Structure across Basins in Lake Erie, a Large Freshwater Ecosystem. *Canadian Journal of Fisheries and Aquatic Sciences* **2025**, 82. https://doi.org/10.1139/cjfas-2024-0028.

(6) Xia, Z.; Idowu, I.; Kerr, E.; Klaassen, N.; Assi, H.; Bray, H.; Marvin, C.; Thomas, P. J.; Stetefeld, J.; Tomy, G. T. New Approaches to Reduce Sample Processing Times for the Determination of Polycyclic Aromatic Compounds in Environmental Samples. *Chemosphere* **2021**, 274. https://doi.org/10.1016/j.chemosphere.2021.129738.

(7) Vitharana, N. N.; Halldorson, T.; Xia, Z.; Francisco, O.; Marvin, C.; Thomas, P. J.; Liebzeit, E.; Lucas, A. M.; Moradi, V.; Tomy, G. T. A Validated Approach for Analysis of Heterocyclic Aromatic Compounds in Sediment Samples. *J. Chromatogr. A* **2024**, 1718. https://doi.org/10.1016/j.chroma.2024.464723.

(8) Jensen, D.; Hofler, F.; Ezzell, J.; Richter, B. Rapid Preparation of Environmental Samples by Accelerated Solvent Extraction (ASE). *Polycycl. Aromat. Compd.* **1996**, 9 (1–4), 233–240. https://doi.org/10.1080/10406639608031223.

(9) Hawthorne, S. B.; Grabanski, C. B.; Martin, E.; Miller, D. J. Comparisons of Soxhlet Extraction, Pressurized Liquid Extraction, Supercritical Fluid Extraction and Subcritical Water Extraction for Environmental Solids: Recovery, Selectivity and Effects on Sample Matrix; **2000**; Vol. 892. www.elsevier.com/locate/chroma.

(10) Josefsson, S.; Westbom, R.; Mathiasson, L.; Björklund, E. Evaluation of PLE Exhaustiveness for the Extraction of PCBs from Sediments and the Influence of Sediment

Characteristics. *Anal. Chim. Acta* **2006**, 560 (1–2), 94–102. https://doi.org/10.1016/j.aca.2005.12.044.

(11) Ramos, L.; Vreuls, J.; Brinkman, Uat. Miniaturised Pressurised Liquid Extraction of Polycyclic Aromatic Hydrocarbons from Soil and Sediment with Subsequent Large-Volume Injection-Gas Chromatography; **2000**; Vol. 891. www.elsevier.com/locate/chroma.

(12) Schubert, P.; Fernandez-Escobar, I.; Rosenberg, E.; Bayona, J.-M. Evaluation of Desulfurization Procedures for the Elimination of Sulfur Interferences in the Organotin Analysis of Sediments; **1998**; Vol. 810.

(13) Lee', B. I.; Rives, J. P. Dispersion of Alumina Powders in Nonaqueous Media. *Colloids and Surfaces* **1991**, *56*, 25–43.

(14) Donald Dugger, by L.; Stanton, J. H.; Irby, B. N.; McConnell, B. L.; Cummings, W. W.; Maatman, R. W. Exchange of Metal Ions with Acidic Silanol Groups The Exchange of Twenty Metal Ions with the Weakly Acidic Silanol Group of Silica Gel12; **1954**; Vol. 173. https://pubs.acs.org/sharingguidelines.

(15) Vitharana, N. N.; Halldorson, T.; Dereviankin, M.; Paddock, B.; Gerring, A.; Marvin, C.; Helm, P.; Thomas, P. J.; Idowu, I. G.; Tomy, G. T. Heterocyclic Aromatic Compounds in Lake Ontario Sediments: Spatiotemporal Distribution and New Approaches to Fingerprinting. *Science of the Total Environment* **2025**, 986. https://doi.org/10.1016/j.scitotenv.2025.179802.

(16) Griffitt, K. R.; Craun, J. C. Gel Permeation Chromatographic System: An Evaluation. *Journal of Association of Official Analytical Chemists* **1974**, 57 (1), 168–172. https://doi.org/10.1093/jaoac/57.1.168.

(17) Idowu, I.; Francisco, O.; Thomas, P. J.; Johnson, W.; Marvin, C.; Stetefeld, J.; Tomy, G. T. Validation of a Simultaneous Method for Determining Polycyclic Aromatic Compounds and Alkylated Isomers in Biota. *Rapid Communications in Mass Spectrometry* **2018**, 32 (3), 277–287. https://doi.org/10.1002/rcm.8035.

(18) Hu, T.; Xia, Y.; Wang, Y.; Lin, L.; An, R.; Xu, L.; Qing, X. Fast Determination and Source Apportionment of Eight Polycyclic Aromatic Hydrocarbons in PM10 Using the Chemometric-Assisted HPLC-DAD Method. *Chemosensors* **2024**, 12 (10). https://doi.org/10.3390/chemosensors12100220.

(19) Hill, H. H.; Chan, K. W.; Karasek', F. W. Extraction of organic compounds from .urborne particulate matter for gas chromatographic analysis; Elseviet Scientific Publishing Company, **1977**; Vol. 13.

(20) Guo, X.; Zhang, W.; Gu, J.; Chen, F.; Yang, Q. The Determination of the Level, Source, and Risk of Polycyclic Aromatic Hydrocarbon Content in Traditional Chinese Medicines Using a QuEChERS Based Extraction and HPLC-UV-FLD *Analysis. J. Liq. Chromatogr. Relat. Technol*. **2021**, 44 (3–4), 210–219. https://doi.org/10.1080/10826076.2021.1883649.

(21) Kicinski, H. G.; Adamek, S.; Kettrup, A. Trace enrichment and hplc analysis of polycyclic aromatic-hydrocarbons in environmental-samples, using solid-phase extraction in connection with uv/vis diode-array and fluorescence detection. *Chromatographia* **1989**, 28 (3–4), 203–208. https://doi.org/10.1007/BF02319648.

(22) Yang, L.; Wu, J.; Zheng, M.; Cao, Z.; Li, C.; Shi, M.; Liu, G. Non-Target Screening of Organic Pollutants and Target Analysis of Halogenated Polycyclic Aromatic Hydrocarbons in the Atmosphere around Metallurgical Plants by High-Resolution GC/Q-TOF-MS. *Environ. Sci. Eur.* **2020**, 32 (1). https://doi.org/10.1186/s12302-020-00376-9.

(23) Achten, C.; Marin-Enriquez, O.; Behrends, B.; Kupich, S.; Lutter, A.; Korth, R.; Andersson, J. T. Polycyclic Aromatic Compounds Including Non-Target and 71 Target Polycyclic Aromatic Hydrocarbons in Scrubber Discharge Water and Their Environmental Impact. *Mar. Pollut. Bull.* **2024**, 208. https://doi.org/10.1016/j.marpolbul.2024.116790.

(24) Ahad, J. M. E.; Macdonald, R. W.; Parrott, J. L.; Yang, Z.; Zhang, Y.; Siddique, T.; Kuznetsova, A.; Rauert, C.; Galarneau, E.; Studabaker, W. B.; Evans, M.; McMaster, M. E.; Shang, D. Polycyclic Aromatic Compounds (PACs) in the Canadian Environment: A Review of Sampling Techniques, Strategies and Instrumentation. *Environmental Pollution*. Elsevier Ltd November 1, **2020**. https://doi.org/10.1016/j.envpol.2020.114988.

(25) Noro, K.; Omagari, R.; Ito, K.; Wang, Q.; Sei, K.; Miyake, Y.; Amagai, T. Sampling, Pretreatment, Instrumental Analysis, and Observed Concentrations of Polycyclic Aromatic Hydrocarbons, Polychlorinated Naphthalenes, and Halogenated Polycyclic Aromatic Hydrocarbons: A Review. *TrAC - Trends in Analytical Chemistry*. Elsevier B.V. December 1, **2023**. https://doi.org/10.1016/j.trac.2023.117384.

(26) Jin, R.; Liu, G.; Zhou, X.; Zhang, Z.; Lin, B.; Liu, Y.; Qi, Z.; Zheng, M. Analysis of Polycyclic Aromatic Hydrocarbon Derivatives in Environment. *TrAC - Trends in Analytical Chemistry*. Elsevier B.V. March 1, **2023**. https://doi.org/10.1016/j.trac.2023.116942.

(27) Martin, A. J. P.; Synge, R. L. M. I5I. A new form of chromatogram employing two liquid phases a theory of chromatography 2. Application to the micro-determination of the higher monoamino-acids in proteins. http://portlandpress.com/biochemj/article-pdf/35/12/1358/795417/bj0351358.pdf.

(28) James, A. T.; Martin, A. J. P. Liquid-gas partition chromatography. *Biochemical Journal* **1951**, 48 (5), R7–R7.

(29) James, A. T.; Martin, A. J. P. Gas-Liquid Partition Chromatography: The Separation and Micro-Estimation of Volatile Fatty Acids from Formic Acid to. Dodecanoic Acid; Vol. 50. http://portlandpress.com/biochemj/article-pdf/50/5/679/800348/bj0500679.pdf.

(30) Thomson, J. J. XXVI. Rays of Positive Electricity. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **1911**, 21 (122), 225–249. https://doi.org/10.1080/14786440208637024.

(31) Flenker, U. Isotope Ratio Mass Spectrometry - History and Terminology in Brief. *Drug Testing and Analysis*. John Wiley and Sons Ltd **2012**, pp 893–896. https://doi.org/10.1002/dta.1399.

(32) Zarrouk, E.; Lenski, M.; Bruno, C.; Thibert, V.; Contreras, P.; Privat, K.; Ameline, A.; Fabresse, N. High-Resolution Mass Spectrometry: Theoretical and Technological Aspects. *Toxicologie Analytique et Clinique* **2022**, 34 (1, Supplement), 3–18. https://doi.org/https://doi.org/10.1016/j.toxac.2021.11.002.

(33) Zubarev, R. A.; Makarov, A. Orbitrap Mass Spectrometry. *Anal. Chem.* **2013**, 85 (11), 5288–5296. https://doi.org/10.1021/ac4001223.

(34) Zoccali, M.; Tranchida, P. Q.; Mondello, L. Fast Gas Chromatography-Mass Spectrometry: A Review of the Last Decade. *TrAC - Trends in Analytical Chemistry.* Elsevier B.V. September 1, **2019**, pp 444–452. https://doi.org/10.1016/j.trac.2019.06.006.

(35) Medhe, S. Ionization Techniques in Mass Spectrometry: A Review. *Mass Spectrom. Purif. Tech.* **2018**, 04 (01). https://doi.org/10.4172/2469-9861.1000126.

(36) Krauss, M.; Singer, H.; Hollender, J. LC-High Resolution MS in Environmental Analysis: From Target Screening to the Identification of Unknowns. *Anal. Bioanal. Chem.* **2010**, 397 (3), 943–951. https://doi.org/10.1007/s00216-010-3608-9.

(37) Peacock, P. M.; Zhang, W. J.; Trimpin, S. Advances in Ionization for Mass Spectrometry. *Analytical Chemistry*. American Chemical Society January 3, **2017**, pp 372–388. https://doi.org/10.1021/acs.analchem.6b04348.

(38) Trenholm, R. A.; Vanderford, B. J.; Holady, J. C.; Rexing, D. J.; Snyder, S. A. Broad Range Analysis of Endocrine Disruptors and Pharmaceuticals Using Gas Chromatography and Liquid Chromatography Tandem Mass Spectrometry. *Chemosphere* **2006**, 65 (11), 1990–1998. https://doi.org/10.1016/j.chemosphere.2006.07.004.

(39) Skoczyńska, E.; Korytár, P.; De Boer, J. Maximizing Chromatographic Information from Environmental Extracts by GCxGC-ToF-MS. *Environ. Sci. Technol*. **2008**, 42 (17), 6611–6618. https://doi.org/10.1021/es703229t.

(40) Jonscher, K. R.; Yates, J. R. Review: The Quadrupole Ion Trap Mass Spectrometer-A Small Solution to a Big Challenge; **1997**; Vol. 244.

(41) Haag, A. M. Mass Analyzers and Mass Spectrometers. In Modern Proteomics – Sample Preparation, Analysis and Practical Applications; Mirzaei, H., Carrasco, M., Eds.; Springer International Publishing: Cham, **2016**; pp 157–169. https://doi.org/10.1007/978-3-319-41448-5_7.

(42) Petrović, M.; Hernando, M. D.; Díaz-Cruz, M. S.; Barceló, D. Liquid Chromatography-Tandem Mass Spectrometry for the Analysis of Pharmaceutical Residues in Environmental Samples: A Review. *Journal of Chromatography A.* March 4, **2005**, pp 1–14. https://doi.org/10.1016/j.chroma.2004.10.110.

(43) Santos, F. J.; Galceran, M. T. The Application of Gas Chromatography to Environmental Analysis; **2002**.

(44) Richardson, S. D. Mass Spectrometry in Environmental Sciences. *Chemical Reviews*. February **2001**, pp 211–254. https://doi.org/10.1021/cr990090u.

(45) Richardson, S. D. Environmental Mass Spectrometry: Emerging Contaminants and Current Issues. *Analytical Chemistry.* January 17, **2012**, pp 747–778. https://doi.org/10.1021/ac202903d.

(46) Machado, M. E.; Nascimento, M. M.; Bomfim Bahia, P. V.; Martinez, S. T.; Bittencourt de Andrade, J. Analytical Advances and Challenges for the Determination of Heterocyclic Aromatic Compounds (NSO-HET) in Sediment: A Review. *TrAC - Trends in Analytical Chemistry.* Elsevier B.V. May 1, **2022**. https://doi.org/10.1016/j.trac.2022.116586.

(47) Santos, F. J.; Galceran, M. T. Modern Developments in Gas Chromatography-Mass Spectrometry-Based Environmental Analysis. *Journal of Chromatography A*. Elsevier June 6, **2003**, pp 125–151. https://doi.org/10.1016/S0021-9673(03)00305-4.

(48) Gross, J. H. Tandem Mass Spectrometry. In Mass Spectrometry: A Textbook; Gross, J. H., Ed.; Springer International Publishing: Cham, **2017**; pp 539–612. https://doi.org/10.1007/978-3-319-54398-7_9.

(49) De Hoffmann, E. Tandem Mass Spectrometry: A Primer. *Journal of Mass Spectrometry*. February **1996**, pp 129–137. https://doi.org/10.1002/(SICI)1096-9888(199602)31:2<129::AID-JMS305>3.0.CO;2-T.

(50) Marriott, P. J.; Haglund, P.; Ong, R. C. Y. A Review of Environmental Toxicant Analysis by Using Multidimensional Gas Chromatography and Comprehensive GC. *Clinica Chimica Acta* **2003** 328, pp 1-19 www.elsevier.com/locate/clinchim.

(51) Johnson, W.; Idowu, I.; Francisco, O.; Marvin, C.; Thomas, P. J.; Stetefeld, J.; Tomy, G. T. Enumeration of the Constitutional Isomers of Environmentally Relevant Substituted Polycyclic Aromatic Compounds. *Chemosphere* **2018**, 202, 9–16. https://doi.org/10.1016/j.chemosphere.2018.03.035.

(52) Muscalu, A. M.; Górecki, T. Comprehensive Two-Dimensional Gas Chromatography in Environmental Analysis. *TrAC - Trends in Analytical Chemistry.* Elsevier B.V. September 1, **2018**, pp 225–245. https://doi.org/10.1016/j.trac.2018.07.001.

(53) Pani, O.; Górecki, T. Comprehensive Two-Dimensional Gas Chromatography (GC×GC) in Environmental Analysis and Monitoring. *Analytical and Bioanalytical Chemistry*. Springer Verlag **2006**, pp 1013–1023. https://doi.org/10.1007/s00216-006-0568-1.

(54) Dallüge, J.; Beens, J.; Brinkman, U. A. T. Comprehensive Two-Dimensional Gas Chromatography: A Powerful and Versatile Analytical Tool. *Journal of Chromatography A*. Elsevier June 6, **2003**, pp 69–108. https://doi.org/10.1016/S0021-9673(03)00242-5.

(55) Idowu, I.; Johnson, W.; Francisco, O.; Obal, T.; Marvin, C.; Thomas, P. J.; Sandau, C. D.; Stetefeld, J.; Tomy, G. T. Comprehensive Two-Dimensional Gas Chromatography High-Resolution Mass Spectrometry for the Analysis of Substituted and Unsubstituted Polycyclic Aromatic Compounds in Environmental Samples. *J. Chromatogr. A* **2018**, 1579, 106–114. https://doi.org/10.1016/j.chroma.2018.10.030.

(56) Marshall, A. G.; Hendrickson, C. L. High-Resolution Mass Spectrometers. *Annual Review of Analytical Chemistry.* **2008**, pp 579–599. https://doi.org/10.1146/annurev.anchem.1.031207.112945.

(57) Hernández, F.; Sancho, J. V.; Ibáñez, M.; Abad, E.; Portolés, T.; Mattioli, L. Current Use of High-Resolution Mass Spectrometry in the Environmental Sciences. *Analytical and Bioanalytical Chemistry.* May **2012**, pp 1251–1264. https://doi.org/10.1007/s00216-012-5844-7.

(58) Canchola, A.; Tran, L. N.; Woo, W.; Tian, L.; Lin, Y. H.; Chou, W. C. Advancing Non-Target Analysis of Emerging Environmental Contaminants with Machine Learning: Current Status and Future Implications. *Environment International.* Elsevier Ltd April 1, **2025**. https://doi.org/10.1016/j.envint.2025.109404.

(59) Mamyrin, B. A. Time-of-Flight Mass Spectrometry (Concepts, Achievements, and Prospects); **2001**.

(60) Boesl, U. Time-of-Flight Mass Spectrometry: Introduction to the Basics. *Mass Spectrometry Reviews*. John Wiley and Sons Inc. January 1, **2017**, pp 86–109. https://doi.org/10.1002/mas.21520.

(61) Guilhaus, M.; Selby, D.; Mlynski, V. Orthogonal Acceleration Time-of-Flight Mass Spectrometry. *Mass Spectrom. Rev.* **2000**, 19 (2), 65–107. https://doi.org/10.1002/(SICI)1098-2787(2000)19:2<65::AID-MAS1>3.0.CO;2-E.

(62) Weickhardt, C.; Moritz, F.; Grotemeyer, J. Time-of-Flight Mass Spectrometry: State-of-the-Art in Chemical Analysis and Molecular Science. *Mass Spectrom. Rev*. **1996**, 15 (3), 139–162. https://doi.org/10.1002/(SICI)1098-2787(1996)15:3<139::AID-MAS1>3.0.CO;2-J.

(63) Itoh, N.; Numata, M.; Aoyagi, Y.; Yarita, T. Accurate Quantification of Polycyclic Aromatic Hydrocarbons in Environmental Samples Using Deuterium-Labeled Compounds as Internal Standards. *Analytical Sciences* **2008**, 24 (9), 1193–1197. https://doi.org/10.2116/analsci.24.1193.

(64) Idowu, I. G.; Xia, Z.; Sandau, C. D.; Misselwitz, M.; Thomas, P.; Marvin, C.; Tomy, G. T. Comparison of Different Approaches to Quantify Substituted Polycyclic Aromatic Compounds. *J. Chromatogr. A* **2021**, 1651. https://doi.org/10.1016/j.chroma.2021.462317.

(65) Otto, M. Chemometrics Statistics and Computer Application in Analytical Chemistry, Fourth.; WILEY-VCH, **2024**.

(66) Kumar, R.; Sharma, V. Chemometrics in Forensic Science. *TrAC - Trends in Analytical Chemistry.* Elsevier B.V. August 1, **2018**, pp 191–201. https://doi.org/10.1016/j.trac.2018.05.010.

(67) Miller, J. N. Basic Statistical Methods for Analytical Chemistry Part 2. Calibration and Regression Methods* A Review; **1991**; Vol. 116.

(68) Brüggemann, L.; Morgenstern, P.; Wennrich, R. Comparison of Regression Techniques for Linear Calibration. Accreditation and Quality Assurance **2005**, 10 (7), 344–351. https://doi.org/10.1007/s00769-005-0936-0.

(69) Chen, T.; Martin, E. Bayesian Linear Regression and Variable Selection for Spectroscopic Calibration. *Anal. Chim. Acta* **2009**, 631 (1), 13–21. https://doi.org/10.1016/j.aca.2008.10.014.

(70) Aboal-Somoza, M.; Crujeiras, R. M. Misuse of Linear Regression Technique in Analytical Chemistry? *J. Chem. Educ.* **2024**, 101 (3), 1062–1070. https://doi.org/10.1021/acs.jchemed.3c01042.

(71) Chen, H. Y.; Chen, C. Evaluation of Calibration Equations by Using Regression Analysis: An Example of Chemical Analysis. *Sensors* **2022**, 22 (2). https://doi.org/10.3390/s22020447.

(72) Raposo, F.; Barceló, D. Assessment of Goodness-of-Fit for the Main Analytical Calibration Models: Guidelines and Case Studies. *TrAC - Trends in Analytical Chemistry.* Elsevier B.V. October 1, **2021**. https://doi.org/10.1016/j.trac.2021.116373.

(73) Van Loco, J.; Elskens, M.; Croux, C.; Beernaert, H. Linearity of Calibration Curves: Use and Misuse of the Correlation Coefficient. *Accreditation and Quality Assurance* **2002**, 7 (7), 281–285. https://doi.org/10.1007/s00769-002-0487-6.

(74) Almeida, ˜ A M; Castel-Branco, M. M.; Falcao, A. C. Linear Regression for Calibration Lines Revisited: Weighting Schemes for Bioanalytical Methods; **2002**; Vol. 774. www.elsevier.com/locate/chromb.

(75) Habibzadeh, F. Data Distribution: Normal or Abnormal? *J. Korean Med. Sci*. **2024**, 39 (3). https://doi.org/10.3346/jkms.2024.39.e35.

(76) Ahsanullah, M.; Kibria, B. M. G.; Shakil, M. Normal Distribution. In Normal and Student´s t Distributions and Their Applications M., Eds.; Atlantis Press: Paris, **2014**; pp 7–50. https://doi.org/10.2991/978-94-6239-061-4_2.

(77) Reimann, C.; Filzmoser, P. Normal and Lognormal Data Distribution in Geochemistry Death of a Myth. Consequences for the Statistical Treatment of Geochemical and Environmental Data. *Environmental Geology* **1999**, 39.

(78) Hatem, G.; Zeidan, J.; Goossens, M.; Moreira, C. Normality testing methods and the importance of skewness and kurtosis in statistical analysis. *BAU Journal - Science and Technology* **2022**, 3 (2). https://doi.org/10.54729/ktpe9512.

(79) Kilgour, B. W.; Somers, K. M.; Barrett, T. J.; Munkittrick, K. R.; Francis, A. P. Testing against “Normal” with Environmental Data. Integr. *Environ. Assess. Manag*. **2017**, 13 (1), 188–197. https://doi.org/10.1002/ieam.1775.

(80) Tebby, C.; Joachim, S.; Van den Brink, P. J.; Porcher, J. M.; Beaudouin, R. Analysis of Community-Level Mesocosm Data Based on Ecologically Meaningful Dissimilarity Measures and Data Transformation. *Environ. Toxicol. Chem*. **2017**, 36 (6), 1667–1679. https://doi.org/10.1002/etc.3701.

(81) Yazici, B.; Yolacan, S. A Comparison of Various Tests of Normality. J. *Stat. Comput. Simul*. **2007**, 77 (2), 175–183. https://doi.org/10.1080/10629360600678310.

(82) González-Estrada, E.; Villaseñor, J. A.; Acosta-Pech, R. Shapiro-Wilk Test for Multivariate Skew-Normality. *Comput. Stat*. **2022**, 37 (4), 1985–2001. https://doi.org/10.1007/s00180-021-01188-y.

(83) Andersson, A. Mechanisms for Log Normal Concentration Distributions in the Environment. *Sci. Rep*. **2021**, 11 (1). https://doi.org/10.1038/s41598-021-96010-6.

(84) Legendre, P.; Gallagher, E. D. Ecologically Meaningful Transformations for Ordination of Species Data. *Oecologia* **2001**, 129 (2), 271–280. https://doi.org/10.1007/s004420100716.

(85) Legendre, P.; Borcard, D. Box–Cox-Chord Transformations for Community Composition Data Prior to Beta Diversity Analysis. *Ecography* **2018**, 41 (11), 1820–1824. https://doi.org/10.1111/ecog.03498.

(86) Likando, N. M.; Chipandwe, M. S. Statistical Investigation of Climate and Landfill Age Impacts on Kupferberg Landfill Leachate Composition: One-Way ANOVA Analysis. *Discover Water* **2024**, 4 (1). https://doi.org/10.1007/s43832-024-00102-z.

(87) Potvin, C. ANOVA: Experiments in Controlled Environments. *Design and Analysis of Ecological Experiments* **1998**

(88) Nanda, A.; Mohapatra, Dr. B. B.; Mahapatra, A. P. K.; Mahapatra, A. P. K.; Mahapatra, A. P. K. Multiple Comparison Test by Tukey's Honestly Significant Difference (HSD): Do the Confident Level Control Type I Error. *International Journal of Statistics and Applied Mathematics* **2021**, 6 (1), 59–65. https://doi.org/10.22271/maths.2021.v6.i1a.636.

(89) Wang, Y.; Li, D.; Wei, P. Powerful Tukey's One Degree-of-Freedom Test for Detecting Gene-Gene and Gene-Environment Interactions. *Cancer Inform.* **2015**, 14, 209–218. https://doi.org/10.4137/CIN.S17305.

(90) Jan, S. L.; Shieh, G. Sample Size Determinations for Welch's Test in One-Way Heteroscedastic ANOVA. *British Journal of Mathematical and Statistical Psychology* **2014**, 67 (1), 72–93. https://doi.org/10.1111/BMSP.12006;REQUESTEDJOURNAL:JOURNAL:20448317;WGROUP:STRING:PUBLICATION.

(91) Alekseyenko, A. V. Multivariate Welch T-Test on Distances. *Bioinformatics* **2016**, 32 (23), 3552–3558. https://doi.org/10.1093/BIOINFORMATICS/BTW524.

(92) The Fitness for Purpose of Analytical Methods : A Laboratory Guide to Method Validation and Related Topics, Second edition, **2014**.; Magnusson, B., Örnemark, U., Eds.

(93) Dutilleul, P.; Stockwell, J. D.; Frigon, D.; Legendre, P. The Mantel Test versus Pearson's Correlation Analysis: Assessment of the Differences for Biological and Environmental Studies; **2000**; Vol. 5. https://about.jstor.org/terms.

(94) Reimann, C.; Filzmoser, P.; Hron, K.; Kynčlová, P.; Garrett, R. G. A New Method for Correlation Analysis of Compositional (Environmental) Data – a Worked Example. *Science of the Total Environment* **2017**, 607–608, 965–971. https://doi.org/10.1016/j.scitotenv.2017.06.063.

(95) Xiao, C.; Ye, J.; Esteves, R. M.; Rong, C. Using Spearman's Correlation Coefficients for Exploratory Data Analysis on Big Dataset. In Concurrency and Computation: Practice and Experience; John Wiley and Sons Ltd, **2016**; Vol. 28, pp 3866–3878. https://doi.org/10.1002/cpe.3745.

(96) Rajesh, G.; Chaturvedi, A. Correlation Analysis and Statistical Characterization of Heterogeneous Sensor Data in Environmental Sensor Networks. *Computer Networks* **2019**, 164. https://doi.org/10.1016/j.comnet.2019.106902.

(97) Filzmoser, P.; Hron, K.; Reimann, C. Univariate Statistical Analysis of Environmental (Compositional) Data: Problems and Possibilities. *Science of the Total Environment* **2009**, 407 (23), 6100–6108. https://doi.org/10.1016/j.scitotenv.2009.08.008.

(98) Lazo, P.; Qarri, F.; Allajbeu, S.; Kane, S.; Bekteshi, L.; Frontasyeva, M.; Stafilov, T. Multivariate Analysis. In The Evaluation of Air Quality in Albania by Moss Biomonitoring and Metals Atmospheric Deposition; *T., Eds*.; Springer International Publishing: Cham, **2021**; pp 89–96. https://doi.org/10.1007/978-3-030-62355-5_8.

(99) Bro, R.; Smilde, A. K. Principal Component Analysis. Analytical Methods. Royal Society of Chemistry May 7, **2014**, pp 2812–2831. https://doi.org/10.1039/c3ay41907j.

(100) Zitko, V. Principal Component Analysis in the Evaluation of Environmental Data; **1994**; Vol. 28.

(101) Armstrong, G.; Martino, C.; Rahman, G.; Gonzalez, A.; Vázquez-Baeza, Y.; Mishne, G.; Knight, R. Uniform Manifold Approximation and Projection (UMAP) Reveals Composite Patterns and Resolves Visualization Artifacts in Microbiome Data. mSystems **2021**, 6 (5). https://doi.org/10.1128/msystems.00691-21.

(102) Healy, J.; McInnes, L. Uniform Manifold Approximation and Projection. *Nature Reviews* Methods Primers **2024**, 4 (1), 82. https://doi.org/10.1038/s43586-024-00363-x.

(103) Capblancq, T.; Luu, K.; Blum, M. G. B.; Bazin, E. Evaluation of Redundancy Analysis to Identify Signatures of Local Adaptation. *Mol. Ecol. Resour.* **2018**, 18 (6), 1223–1233. https://doi.org/10.1111/1755-0998.12906.

(104) Sadyś, M.; Strzelczak, A.; Grinn-Gofroń, A.; Kennedy, R. Application of Redundancy Analysis for Aerobiological Data. Int. J. *Biometeorol*. **2015**, 59 (1), 25–36. https://doi.org/10.1007/s00484-014-0818-4.

(105) Malley, C. S.; Braban, C. F.; Heal, M. R. The Application of Hierarchical Cluster Analysis and Non-Negative Matrix Factorization to European Atmospheric Monitoring Site Classification. *Atmos. Res*. **2014**, 138, 30–40. https://doi.org/10.1016/j.atmosres.2013.10.019.

(106) Saraçli, S.; Do, N.; Gan, ˘; Ismet, ˙; Gan, D. ˘. Comparison of Hierarchical Cluster Analysis Methods by Cophenetic Correlation; **2013**. http://www.journalofinequalitiesandapplications.com/content/2013/1/203.

(107) Xu, C.; Gao, L.; Zheng, M.; Qiao, L.; Wang, K.; Huang, D.; Wang, S. Nontarget Screening of Polycyclic Aromatic Compounds in Atmospheric Particulate Matter Using Ultrahigh Resolution Mass Spectrometry and Comprehensive Two-Dimensional Gas Chromatography. *Environ. Sci. Technol*. **2021**, 55 (1), 109–119. https://doi.org/10.1021/acs.est.0c02290.

(108) Luo, R.; Schrader, W. Getting a Better Overview of a Highly PAH Contaminated Soil: A Non-Targeted Approach Assessing the Real Environmental Contamination. *J. Hazard. Mater*. **2021**, 418. https://doi.org/10.1016/j.jhazmat.2021.126352.

(109) Veenaas, C.; Bignert, A.; Liljelind, P.; Haglund, P. Nontarget Screening and Time-Trend Analysis of Sewage Sludge Contaminants via Two-Dimensional Gas Chromatography-High Resolution Mass Spectrometry. *Environ. Sci. Technol*. **2018**, 52 (14), 7813–7822. https://doi.org/10.1021/acs.est.8b01126.

(110) Titaley, I. A.; Lam, M. M.; Bülow, R.; Enell, A.; Wiberg, K.; Larsson, M. Characterization of Polycyclic Aromatic Compounds in Historically Contaminated Soil by Targeted and Non-Targeted Chemical Analysis Combined with in Vitro Bioassay. *Environmental Pollution* **2021**, 289. https://doi.org/10.1016/j.envpol.2021.117910.

(111) Luo, R.; Schrader, W. Development of a Non-Targeted Method to Study Petroleum Polyaromatic Hydrocarbons in Soil by Ultrahigh Resolution Mass Spectrometry Using Multiple Ionization Methods. *Polycycl. Aromat. Compd.* **2022**, 42 (2), 643–658. https://doi.org/10.1080/10406638.2020.1748665.

(112) Schemeth, D.; Nielsen, N. J.; Andersson, J. T.; Christensen, J. H. A Tiered Analytical Approach for Target, Non-Target and Suspect Screening Analysis of Polar Transformation Products of Polycyclic Aromatic Compounds. *Chemosphere* **2019**, 235, 175–184. https://doi.org/10.1016/j.chemosphere.2019.06.149.

# CHAPTER 3 A VALIDATED APPROACH FOR ANALYSIS OF HETEROCYCLIC AROMATIC COMPOUNDS IN SEDIMENT SAMPLES

The chapter presents the targeted analytical methods I developed based on GC-MS/MS for a suite of unsubstituted, alkylated, and halogenated HPACs. The developed MRM methods were used for all HPAC-targeted analyses done throughout my research. This chapter includes all the details on MRM method development and method performance characteristics. In the last section of this chapter, I applied the method to determine concentrations of PACs in sediments from Hamilton Harbour.

A version of this chapter has already been published as Vitharana, Nipuni N; Halldorson, Thor; Xia, Zhe; Francisco, Olga; Marvin, Chris; Thomas, Philippe J; Liebzeit, Erin; Lucas, Amica-Mariae; Moradi, Vida; Tomy, Gregg T; A validated approach for analysis of heterocyclic aromatic compounds in sediment samples, Journal of Chromatography A, 2024, volume 1718, 464723. Copyright © 2024. B.V. Elsevier. Reprinted with permission.

## 3.1 ABSTRACT

The scientific literature is replete with analytical methods for the analysis of homocyclic aromatic compounds, especially polycyclic aromatic hydrocarbons and their alkylated analogs. However, there is a paucity of methods for the analysis of nitrogen-, sulfur-, and oxygen-containing polycyclic aromatic compounds (PACs). The lack of commercially available analytical standards, the presence of many structural derivatives and isomers, and the lack of certified reference materials all contribute to the inherent challenges in measuring these compounds. Gas chromatography coupled with a tandem mass spectrometer was used to develop two multiple reaction monitoring methods to detect and quantify fifty-three non-halogenated and halogenated hetero-polycyclic aromatic compounds (HPACs). Because of their greater polarity, strongly non-polar solvents typically employed to extract homocyclic PACs from sediment samples did not yield acceptable recoveries of our target analytes. By adding ethyl acetate to dichloromethane (50:50), recoveries of our target analytes using accelerated solvent extraction increased markedly. The performance characteristics of the validated method including accuracy [> than 67% for 46 (out of 53) analytes], inter- and intra-day precision [<30% for all analytes, (expressed as relative standard deviation)], limits of detection (0.1 to 2.3 ng/g) and quantitation (1.5 to 7.6 ng/g) imply that the method is fit for its intended purpose. A sediment sample from a known contaminated site in Canada was analyzed for both homo- and hetero-PACs. Measured concentrations of $\Sigma_{27}$HPAC (7.3 μg/g, dry weight) were significantly smaller ($p<0.05$) than $\Sigma_{16}$PAHs (80.9 μg/g, dry weight) and $\Sigma_{30}$Alkylated-PAHs (14.2 μg/g, dry weight). These results suggest that the developed method is an effective and efficient approach for the targeted analysis of HPACs and their halogenated derivatives in sediment samples.

## 3.2 INTRODUCTION

Polycyclic aromatic compounds (PACs) continue to be a major focus of environmental studies due to their negative impacts on human health and the environment. Some PACs are persistent, bioaccumulative and toxic (PBT).[1] Many combustion processes, both anthropogenically and naturally occurring, are primary sources of PACs.[2–4] There is a considerable body of information, including a recent review, on the occurrence and distribution of PACs in environmental compartments of key geographical areas in Canada including the Great Lakes, Alberta Oil Sands Region (AOSR) and the Canadian Arctic has reported.[5–9]

Most studies of PACs have focussed on the 16 polyaromatic hydrocarbons (PAHs) identified as priority pollutants by the United States Environmental Protection Agency (US EPA) and on the Canadian Environmental Protection Act, 1999 as Toxic Substance List.[10] However, other classes of PACs can be PBT chemicals, including hetero-polycyclic aromatic compounds (HPACs). Some HPACs can be more toxic than their PAH counterparts.[11,12] Therefore, many scientists argue that we are in need of an updated list of the priority PACs that includes alkylated polyaromatic hydrocarbons and heterocyclic polyaromatic compounds.[13]

Heterocyclic PACs are a diverse class of compounds originating from both natural and anthropogenic sources.[14] The most common HPACs have nitrogen, sulfur, and/or oxygen as the heteroatoms replacing one or more carbon atoms in the polyaromatic structure. Polycyclic aromatic nitrogen heterocycles (PANHs), polycyclic aromatic sulfur heterocycles (PASHs), and polycyclic aromatic oxygen heterocycles (PAOHs) are referred to collectively as NSO-heterocyclic compounds. While PAHs and alkylated-PAHs (APAHs) constituent the most abundant polycyclic aromatic class of compounds in crude oil products, these HPACs are considered minor

constituents. For example, concentrations of PANH, PASH, and PAOH in crude oil range from 0.1 to 2.0, 0.01-3.2, and <0.1 %, respectively.[15,16] Although the number of studies done on halogenated polycyclic aromatic hydrocarbons is comparatively less than that of PAHs and APAHs, several studies have shown that these compounds are detectable in air,[17] soil,[18] sediments,[19] and biota.[9] Therefore, understanding the environmental distribution of halogenated heterocyclic compounds is as important as their non-halogenated analogues.

There has been less focus on the development of methods for concurrent analyses of HPACs,[20,21] compared to homocyclic PACs. Identification and quantitation of HPACs has been a challenge due to the lack of many commercially available standards, large numbers of structural isomers, and a paucity of certified reference materials (CRMs).[21] Some methods developed for the analysis of PAHs and APAHs have been expanded for the analysis of HPACs, including carbazoles and dibenzothiophenes.[22] Some studies have applied non-targeted analysis using two-dimensional high-resolution time-of-flight mass spectrometry (GCxGC/HRTOF-MS); these studies have mainly focussed on qualitative aspects, rather than generating quantitative information.[23,24]

Soil and sediments are primary receiving environments for many PACs [25,26]. Some of the extraction techniques employed to study HPACs in sediments are mechanical agitation [24,27], Soxhlet extraction,[27] accelerated solvent extraction (ASE)[28] and miniaturized solid-liquid extraction.[29] Nevertheless, these techniques often exhibit limitations by focusing only on one or two types of HPACs or a limited number of HPACs.

We hypothesize, then, that since HPACs have similar origins to homocyclic PACs, soils and sediments may also contain HPACs. To test our hypothesis, we first developed and validated a gas chromatographic–tandem mass spectrometric (GC-MS/MS) method for the analysis of fifty-two

non-halogenated and halogenated heterocyclic aromatic compounds. Data quality objectives were obtained in strict accordance with the Eurachem Guidelines for Analytical Methods.[30] A coal tar-contaminated sediment sample from Randle Reef (Hamilton Harbour, ON, Canada) was then analyzed for our target analytes.

## 3.3 MATERIALS AND METHODS

### 3.3.1 Chemicals

Eighty non-halogenated and halogenated heterocyclic aromatic compounds were purchased from Chiron (AS Stiklestadveien 1, 7041 Trondheim, Norway) and Sigma Aldrich Canada Ltd (Oakville, Ontario, Canada). From these, forty-nine standards were included in our final targeted analysis method, which can be found in Table I 1. All high-purity (Optima grade) organic solvents, silica gel (923 grade, 100 - 200 mesh), alumina (60 - 325 mesh), Ottawa sand and anhydrous sodium sulfate, diatomaceous earth (DE) were purchased from Fisher Chemicals (Ottawa, Canada). Labelled PAH standards were purchased from Chiron and were used as the internal standards. For recovery internal standard (RIS) naphthalene-$d_8$, fluorene-$d_{10}$, phenanthrene-$d_{10}$, fluoranthene-$d_{10}$, pyrene-$d_{10}$, benz[a]anthracene-$d_{12}$, chrysene-$d_{12}$, benzo[b]fluoranthene-$d_{12}$, benzo[a]pyrene-$d_{12}$ and benzo[g,h,i]perylene-$d_{14}$ were used. Labeled anthracene (anthracene-$d_{10}$) was used as the instrument performance internal standard (IPIS).

### 3.2.2 Environment sample

Sediments collected in 2018 from Randle Reef (Hamilton Harbour, ON, Canada), the largest PAH-contaminated sediment site (*ca*. 695,000 $m^3$) in the Canadian Great Lakes, were selected for analyses. These samples were supplied to us by the Canadian Association for Laboratory Accreditation (Ottawa, ON, Canada) in 2023 as part of our laboratory's biannual ISO-17025-2018

proficiency testing exercise. Sediment was collected in 2018 into a 20-plastic pail and then stored at 0-5 °C for 10 months. Excess liquid was then decanted, and the sediment was dried at 40 °C for 6 days and then at 25 °C for a further 22 days. Samples were then ground using a Wiley Mill and homogenized using a V-blender for 20 minutes. The water content of the sample was ~8%.

### 3.2.2 Sample extraction and cleanup

Dried sediment samples were extracted using a Dionex Accelerated Solvent Extractor 350 (ThermoFisher, Mississauga, ON). The one-step layerless extraction for PAHs and APAHs, described in Xia et al.,[31] was modified to extract all three types of HPACs. A fiberglass filter was placed at the bottom of the 34 ml ASE tubes, and a mixture of dispersant was added to the cell. The dispersant mixture contained 5 g of baked silica gel, 4.5 g of 5% deactivated alumina, and 1 g of reduced copper. Following this, 1.00 g of sediment sample was mixed with 1.5 g of baked diatomaceous earth (DE) dispersant and added to the ASE cell. Samples were then spiked with 10.0 µL of a 10 ng/µL mixture of recovery internal standards (RIS) prior to filling the cell with baked Ottawa sand baked at 600 °C for 6 h.

Several extracting solvents, either in combination with others or by themselves, were tested in our study. Compared to dichloromethane (DCM): hexane mixtures and only DCM, a 50:50 mixture of DCM: ethyl acetate gave the most acceptable recoveries for the largest number of our targeted compounds (Table I 2). The ASE extraction parameters were as follows: system pressure 1500 psi, oven temperature 120 °C, static time of 15 min with 2 static cycles, 1:1 DCM: ethyl acetate mixture as the extraction solvent, 60% flush volume and a nitrogen purge for 80 s at 150 psi. Extracts were then dried by passing them through a glass funnel containing sodium sulphate and reduced in volume to 100 µL using a gentle stream of ultra-high purity (UHP) nitrogen. Finally, extracts were

transferred to a GC-vial, spiked with 10.0 µL of 10 ng/µL of instrument performance internal standard (IPIS), and diluted to 1mL final volume with hexane.

### 3.2.3 Gas chromatography tandem mass spectrometry

An Agilent 7890 GC coupled with a triple quadrupole Agilent 7000C mass spectrometer with an electron ionization (EI) source was used for GC-MS/MS analysis. An Agilent J&W DB-5ms ultra inert column (30 m x 0.25 mm, 0.25 µm) and helium were used as the carrier gas at a constant flow rate of 1.2 mL/min. Splitless injections of 1 µL into a glass tapered lined with wool at its base were achieved using a PAL RSI 85 auto sampler into our injector port maintained at 310 °C. The splitless delay time was 0.75 mins. Satisfactory GC separations were obtained using the oven temperature programing used for PAHs and APAHs that was published earlier by our group.[31] Both transfer line and source temperature were at 300 °C [9]. Two separate multiple reaction monitoring (MRM) methods were developed for non-halogenated and halogenated HPACs. The MRM transitions of labelled PAHs were incorporated with the method as internal standards. All MRM transitions and related collision energies are shown in Table 3.1. The two chromatograms for non-halogenated HPACs and halogenated HPACs are depicted in Figures 3.1 and 3.2, respectively.

**Table 3.1:** Analyte, elution time, product ions, MS/MS quantitation and confirmation ions and collision energies (CE 1 and CE 2) used for HPACs and halo-HPACs.

| #STD | STD name | Elution time (min) | Precursor Ion (m/z) | Product Ion 1 (m/z) | CE 1 (eV) | Product Ion 2 (m/z) | CE 2 (eV) |
|---|---|---|---|---|---|---|---|
| **MRM 1 HPACs.** | | | | | | | |
| **Window 1 from 4.0 to 5.1 min (Dwell time = 50 ms)** | | | | | | | |
| ISTD1 | Naphthalene-$d_8$ | 4.90 | 136.0 | 134.0 | 35 | 108.0 | 35 |
| **Window 2 from 5.1 to 6.2 min (Dwell time = 30 ms)** | | | | | | | |
| N1 | Quinoline | 5.30 | 129.1 | 102.0 | 25 | 76.0 | 35 |
| N2 | Isoquinoline | 5.50 | 129.1 | 102.0 | 25 | 76.0 | 35 |
| N3 | Indole | 5.72 | 117.1 | 89.0 | 35 | 90.0 | 5 |
| N4 | 2-Methylquinoline | 5.85 | 143.0 | 115.0 | 25 | 128.0 | 25 |
| S1 | 4-Methylbenzothiophene | 5.90 | 146.9 | 120.9 | 25 | 145.9 | 25 |
| N5 | 3-Methylisoquinoline | 6.00 | 143.0 | 115.0 | 40 | 116.0 | 25 |
| **Window 3 from 6.2 to 7.5 min (Dwell time = 40 ms)** | | | | | | | |
| N6 | 2,5-Dimethyl-1-phenylpyrrole | 6.45 | 171.0 | 170.0 | 15 | 154.0 | 35 |
| S2 | 2,3-Dimethylbenzothiophene | 6.90 | 147.0 | 103.0 | 25 | 115.0 | 25 |
| N7 | 4-Phenylpyridine | 7.20 | 155.0 | 154.0 | 15 | 127.0 | 25 |
| N8 | 2,3-Dimethylindole | 7.35 | 143.9 | 77.0 | 35 | 50.9 | 40 |
| **Window 4 from 7.5 to 9.2 min (Dwell time = 40 ms)** | | | | | | | |
| O1 | Dibenzofuran | 7.60 | 168.0 | 139.0 | 25 | 129.0 | 25 |
| ISTD2 | Fluorene-$d_{10}$ | 8.13 | 176.0 | 174.2 | 45 | 172.1 | 45 |
| S3 | 2,3,4-Trimethylbenzothiophene | 8.20 | 176.0 | 161.0 | 15 | 128.0 | 35 |
| O2 | 2-Methyldibenzofuran | 8.55 | 181.0 | 152.0 | 35 | 151.0 | 40 |
| **Window 5 from 9.2 to 10.2 min (Dwell time = 30 ms)** | | | | | | | |
| S4 | Dibenzothiophene | 9.50 | 184.0 | 182.0 | 25 | 152.0 | 25 |
| ISTD3 | Phenanthrene-$d_{10}$ | 9.70 | 188.0 | 184.1 | 40 | 160.1 | 40 |
| ISTD4 | Anthracene-$d_{10}$ | 9.80 | 188.0 | 184.1 | 40 | 160.1 | 40 |
| N9 | Benzo[h]quinoline | 9.81 | 179.0 | 178.0 | 25 | 151.0 | 35 |
| N10 | 1,2,3,4-Tetrahydrocarbazole | 9.85 | 143.0 | 128.3 | 25 | 115.0 | 25 |
| N11 | Acridine | 9.89 | 179.1 | 178.0 | 25 | 151.0 | 30 |
| N12 | 9-Methylcarbazole | 10.05 | 181.1 | 180.1 | 25 | 152.0 | 35 |
| N13 | Carbazole | 10.12 | 167.1 | 166.1 | 25 | 138.9 | 35 |
| **Window 6 from 10.2 to 11.7 min (Dwell time = 40 ms)** | | | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| S5 | 4-Methyldibenzothiophene | 10.30 | 197.8 | 165.0 | 35 | 152.0 | 40 |
| N14 | 3-Methylbenzo[f]quinoline | 10.60 | 193.0 | 165.0 | 35 | 151.0 | 40 |
| N15 | 2-Methylacridine | 10.85 | 193.0 | 192.0 | 25 | 191.0 | 40 |
| S6 | 2,4-Dimethyldibenzothiophene | 11.45 | 211.8 | 197.0 | 15 | 178.0 | 35 |
| **Window 7 from 11.7 to 14.0 min (Dwell time = 40 ms)** | | | | | | | |
| N16 | 3,6-Dimethylcarbazole | 12.02 | 180.1 | 179.0 | 25 | 152.0 | 25 |
| ISTD5 | Fluoranthene-$d_{10}$ | 12.15 | 212.0 | 210.2 | 40 | 208.1 | 40 |
| S7 | 2,4,7-Trimethyldibenzothiophene | 12.45 | 225.8 | 211.1 | 15 | 178.2 | 35 |
| ISTD6 | Pyrene-$d_{10}$ | 12.73 | 212.0 | 210.2 | 40 | 208.1 | 40 |
| O3 | Benzo[b]naphtho[2,3-d]furan | 13.15 | 188.8 | 187.0 | 35 | 163.0 | 35 |
| N17 | 3-Tert-butyl-9H-carbazole | 13.25 | 208.1 | 180.0 | 15 | 167.0 | 30 |
| **Window 8 from 14.0 to 16.8 min (Dwell time = 50 ms)** | | | | | | | |
| N18 | 9-Phenylcarbazole | 14.90 | 243.0 | 241.0 | 35 | 166.0 | 40 |
| S8 | 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | 16.10 | 209.8 | 208.0 | 25 | 183.9 | 25 |
| N19 | Benz[c]acridine | 16.40 | 229.0 | 228.0 | 25 | 227.0 | 45 |
| **Window 9 from 16.8 to 19.0 min (Dwell time = 50 ms)** | | | | | | | |
| ISTD7 | Chrysene-$d_{12}$ | 17.09 | 240.0 | 236.2 | 30 | 212.2 | 30 |
| ISTD8 | Benz[a]anthracene-$d_{12}$ | 17.25 | 240.0 | 236.2 | 30 | 212.2 | 30 |
| N20 | 11H-Benzo[a]carbazole | 17.41 | 217.1 | 216.0 | 30 | 188.8 | 30 |
| N21 | 7H-benzo[c]carbazole | 18.40 | 217.1 | 216.0 | 25 | 188.8 | 30 |
| S9 | 6-Methylbenzo [b]naphtho[2,1-d]thiophene | 18.50 | 247.8 | 247.0 | 25 | 245.0 | 40 |
| **Window 10 from 19.0 to 25.0 min (Dwell time = 50 ms)** | | | | | | | |
| S10 | 2-Phenyldibenzothiophene | 19.45 | 259.7 | 257.0 | 40 | 256.0 | 40 |
| S11 | 1,3-Diphenylbenzo[c]thiophene | 21.75 | 285.7 | 284.0 | 25 | 252.0 | 40 |
| ISTD9 | Benzo[b]fluranthene-$d_{12}$ | 22.39 | 264.0 | 260.2 | 30 | 236.2 | 30 |
| ISTD10 | Benzo[k]fluranthene-$d_{12}$ | 22.50 | 264.0 | 260.2 | 30 | 236.2 | 30 |
| O4 | Dinaphtho[1,2-b;1',2'-d]furan | 23.40 | 267.8 | 239.0 | 35 | 237.0 | 45 |
| **Window 11 from 25.0 to 32 min (Dwell time = 50 ms)** | | | | | | | |
| N22 | Dibenz[c,h]acridine | 29.16 | 279.0 | 278.0 | 35 | 277.0 | 45 |
| N23 | 13(H)-Dibenzo[a,i]carbazole | 30.41 | 267.0 | 266.0 | 35 | 239.o | 40 |
| ISTD11 | Benzo[ghi]perylene-$d_{14}$ | 31.78 | 288.0 | 288.o | 15 | - | - |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ISTD11 | Benzo[ghi]perylene-$d_{14}$ | 31.78 | 286.2 | 286.2 | 15 | - | - |
| **MRM 2 halo-HPACs** | | | | | | | |
| **Window 1 from 4.0 to 7.2 min (Dwell time = 40 ms)** | | | | | | | |
| ISTD1 | Naphthalene-$d_8$ | 4.92 | 136.0 | 134.0 | 35 | 108.0 | 35 |
| HH1 | 2,3-Dibromothiophene | 4.96 | 241.8 | 161.0 | 15 | 162.8 | 15 |
| HH3 | 2-Chloroquinoline | 6.95 | 162.9 | 128.0 | 15 | 101.0 | 25 |
| HH4 | 2-Bromobenzothiophene | 7.10 | 213.8 | 133.0 | 25 | 135.0 | 25 |
| **Window 2 from 7.2 to 8.5 min (Dwell time = 40 ms)** | | | | | | | |
| HH5 | 2,3-Dichlorobenzothiophene | 7.55 | 201.8 | 166.9 | 15 | 123.0 | 25 |
| HH6 | 2-Bromoquinoline | 7.80 | 127.9 | 101.0 | 15 | 77.0 | 15 |
| HH7 | 5-Chloroindole | 7.90 | 150.9 | 89.0 | 15 | 116.0 | 15 |
| ISTD2 | Fluorene-$d_{10}$ | 8.10 | 176.0 | 174.2 | 45 | 172.1 | 45 |
| HH8 | 3-Bromoisoquinoline | 8.25 | 127.9 | 101.0 | 15 | 77.0 | 25 |
| **Window 3 from 8.5 to 11.7 min (Dwell time = 40 ms)** | | | | | | | |
| HH9 | 5-Bromoindole | 8.77 | 116.0 | 89.0 | 15 | 63.0 | 25 |
| HH10 | 2,3-Dibromobenzothiophene | 9.22 | 131.9 | 92.9 | 25 | 68.9 | 25 |
| ISTD3 | Phenanthrene-$d_{10}$ | 9.70 | 188.0 | 184.1 | 40 | 160.1 | 40 |
| ISTD4 | Anthracene-$d_{10}$ | 9.78 | 188.0 | 184.1 | 40 | 160.1 | 40 |
| HH11 | 9-Chloroacridine | 11.40 | 212.9 | 178.0 | 15 | 151.1 | 25 |
| **Window 4 from 11.7 to 14.0 min (Dwell time = 50 ms)** | | | | | | | |
| ISTD5 | Fluoranthene-$d_{10}$ | 12.12 | 212.0 | 210.2 | 40 | 208.1 | 40 |
| HH12 | 2-Bromodibenzothiophene | 12.20 | 263.7 | 183 | 25 | 185 | 25 |
| HH13 | 9-bromoacridine | 12.42 | 177.9 | 151.1 | 15 | 125.1 | 25 |
| ISTD6 | Pyrene-$d_{10}$ | 12.70 | 212.0 | 210.2 | 40 | 208.1 | 40 |
| **Window 5 from 14.0 to 18.0 min (Dwell time = 50 ms)** | | | | | | | |
| HH14 | 3,6-Dichloro-9h-carbazole | 15.95 | 234.8 | 200.0 | 25 | 164.3 | 35 |
| HH15 | 2,8-Dibromodibenzothiophene | 16.45 | 181.8 | 138.0 | 15 | 137.0 | 35 |
| ISTD7 | Chrysene-$d_{12}$ | 17.19 | 240.0 | 236.2 | 30 | 212.2 | 30 |
| **Window 6 from 18.0 to 32 min (Dwell time = 50 ms)** | | | | | | | |
| HH16 | 3,6-Dibromocarbazole | 19.85 | 324.8 | 243.9 | 15 | 246.2 | 15 |
| ISTD10 | Benzo[k]fluranthene-$d_{12}$ | 22.44 | 264.0 | 260.2 | 30 | 236.2 | 30 |

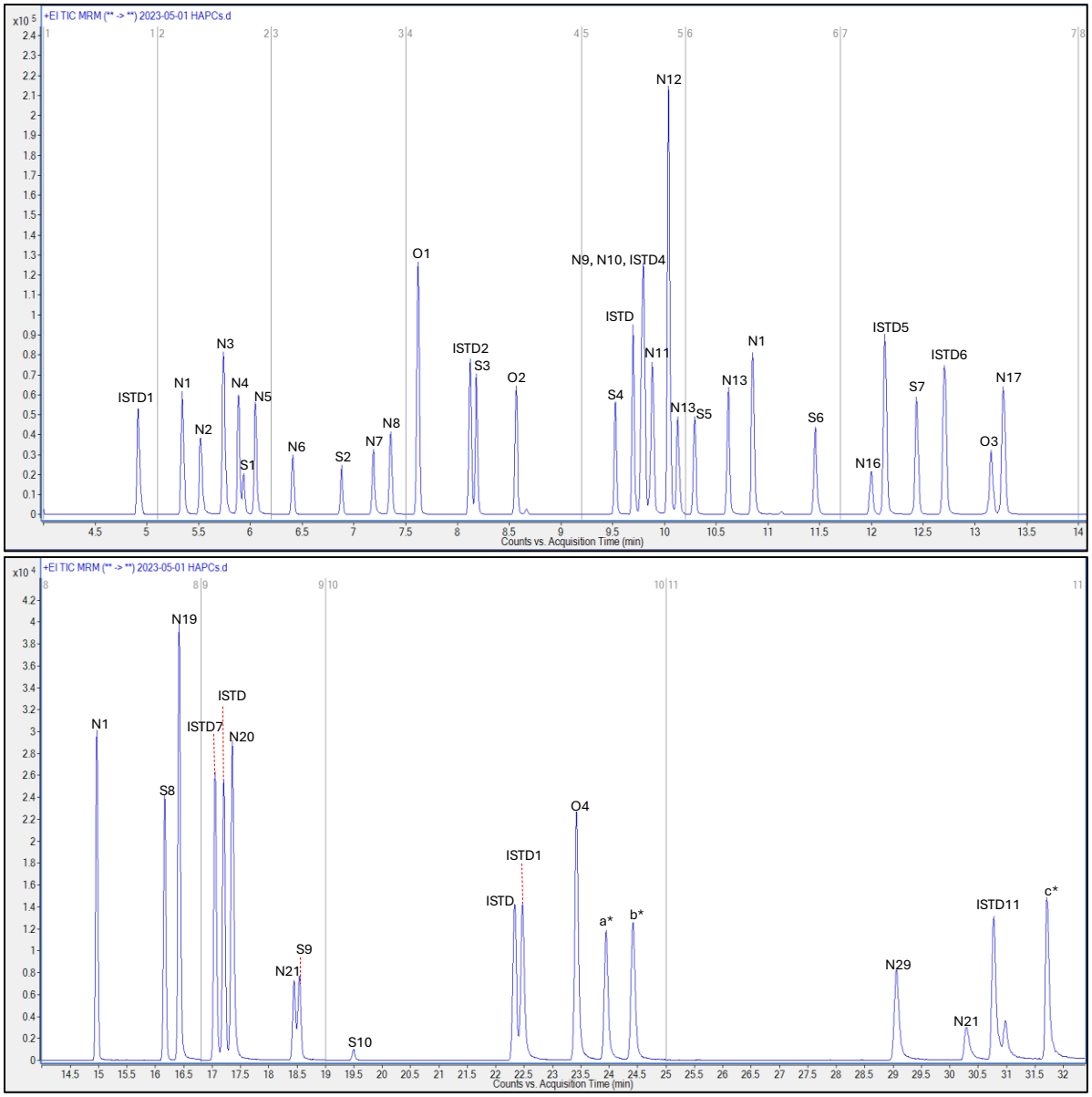



**Figure 3.1:** The MRM ion chromatogram of our HPACs analytical standard (200 pg/µL). Top panel shows analytes that elute from $r_t$: 0-14mins and bottom panel shows analytes that elute from $r_t$: 14-33 min. ISTD4: anthracene-$d_{10}$, $r_t$: 9.76 min, coelutes with N9 and N10 and is not visible in top panel. (Peaks labeled $a^*$, $b^*$ and $c^*$ in bottom panel are mass labeled compounds present in our commercial mixture that were not used in our analytical method.)

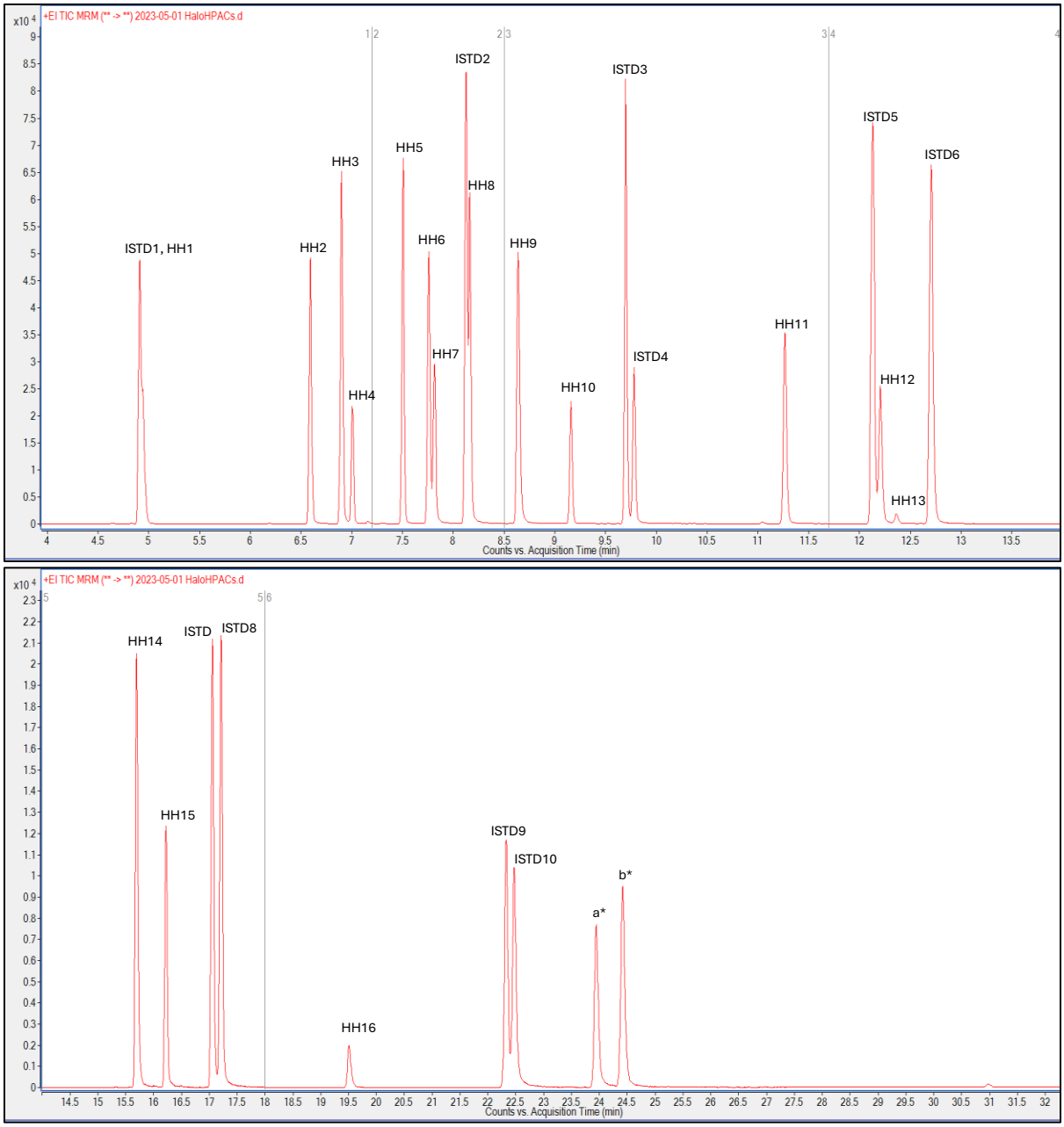



**Figure 3.2:** The MRM ion chromatogram of our halo-HPACs analytical standard (200 pg/µL). Top panel shows analytes that elute from $r_t$: 0-14mins and bottom panel shows analytes that elute from $r_t$: 14-33 min. ISTD1: naphthalene-d8, $r_t$: 4.92 min, coelutes with HH1 and is not visible in the top panel. (Peaks labeled $a^*$ and $b^*$ in bottom panel are mass labeled compounds present in our commercial mixture that were not used in our analytical method.)

### 3.2.4 Method validation

Our method was validated according to the Eurachem Guide -The Fitness for the Purpose of Analytical Methods [30]. The performance characteristics described in the Eurachem guideline are highlighted here as part of our validation. A nine-point (2.5, 5, 10, 25, 50, 100, 200, 400, and 800 pg/µL) calibration curve was developed for all 49 standards using both linear and response factor ($r_f$) models. Three replicates of each standard concentration and three solvent blanks were randomly run to generate our curves. To determine the working range of the method, two mathematical approaches were compared. In the linear model, a traditional calibration curve was constructed by dividing the electronically integrated area of each analyte by an appropriate d-PAH IS and plotting this ratio against the known concentration. The d-PAH that eluted closest to the analyte of interest was used as the IS. For the $r_f$ model, the relative $r_f$ (i.e., corrected for d-PAH) at each concentration level was calculated and then averaged to obtain the average relative $r_f$. The equation below provides the mathematical basis of our approach:

$$Average\ r_f\ =\ \sum\left(\frac{\frac{electronically\ integrated\ area\ of\ analyte}{electronically\ integrated\ area\ of\ d-PAH}}{Concentration\ of\ analyte\ in\ standard}\right)/n$$

where $n$ = the number of standards injected

Because of the stability of the ion signals of our target analytes and RIS between injections, it was unnecessary to use the IPIS in our models. However, the IPIS was used to calculate the recoveries of our RIS, which are reported in Table I 3. The final step in our assessment was to use each model to back-calculate the concentration of each standard. For example, in the linear model, we inserted the ratio of the integrated area of the analyte to that of the d-PAH into the regression equation and calculated a concentration. We then examined the residual plot of the calculated concentration to

the known concentration. This was done for all our analytes. The same approach was adopted for the average $r_f$ model. As an example, the residual plots of quinoline using (i) the linear model and (ii) the average $r_f$ model is shown in Figure 3.3.

Limits of detection- and quantitation (LODs and LOQs) were calculated based on extracting DE dispersant (*n=8*) fortified with 10 pg/ µL of the HPACs mixture. As defined in the Eurachem guidelines, the standard deviations (SD) were calculated using the number of replicates (n=8) and the number of blanks (n=6) for each standard. Limits of detection and LOQ were calculated as 3 × SD and 10 × SD, respectively, and are shown in Table 3.2.

Because of a lack of commercial availability of a standard reference material, we prepared an in-house reference material by fortifying dispersant with our compounds at four different concentrations: 10, 25, 100, and 800 pg/µL. At each concentration, eight replicates were extracted and analyzed using the method described here.

To test the repeatability of the above method, four sets of fortified DE depressant samples (n=8) were extracted. The samples were processed within 24 hours. The relative standard deviations (RSD) were calculated for each concentration level (Table 3.2). To measure the intermediate precision, three sets of 100 pg/µL fortified DE dispersant samples (n=4) were extracted on three separate days. The RSD were calculated for each sample set. Single-factor ANOVA test was performed to statistically analyze the precision within the three days for each compound.

These performance characterization tests were carried out in three sample sets on three consecutive days. In each day, without spiking the analytes of interest, two procedural blanks were prepared and extracted in the same manner as the spiked samples. During the GC-MS/MS analysis, for every 20 sample extracts, one solvent blank was run. Between every GC injection, our syringe was

solvent rinsed with acetone (5 times) and hexane (5 times). Solvent blanks injected after the 800 pg/µL standard confirmed that there was no carryover between injections.

The ruggedness of the method was determined by changing two parameters of the method. One, the mass of each dispersant type (silica gel, 5% deactivated alumina, and reduced copper) was purposely increased by 0.5 g (n=4), and second, the extraction temperature was increased from 120 °C to 140 °C (n=4). The calculated relative standard deviations for each parameter can be found in Table 3.2. For each test parameter of ruggedness one procedural blank was analysed.

**Table 3.2:** Performance characteristics of the HPACs and halo-HPACs method

| # STD | ISTD | % Recovery | | | | Repeatability (RSD%) | | | | Intermediate precision | | LOD (ng/g) | LOQ (ng/g) | Ruggedness % Recovery | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | C1 | C2 | C3 | C4 | C1 | C2 | C3 | C4 | D2 | D3 | | | P1 | P2 |
| **HPACs** | | | | | | | | | | | | | | | |
| N1 | ISTD1 | 114 | 112 | 111 | 97 | 3.1 | 7.7 | 7.8 | 6.9 | 5.4 | 2.1 | 1.3 | 4.3 | 145 | 126 |
| N2 | ISTD1 | 123 | 113 | 108 | 96 | 5.2 | 8.7 | 6.3 | 5.3 | 4.9 | 2.5 | 1.0 | 3.4 | 116 | 78 |
| N3 | ISTD1 | 69 | 32 | 31 | 34 | 10.2 | 16.3 | 13.9 | 29.8 | 13.8 | 4.3 | 1.9 | 6.5 | 83 | 66 |
| N4 | ISTD1 | 110 | 109 | 98 | 94 | 2.4 | 9.3 | 3.3 | 5.3 | 5.6 | 2.4 | 0.8 | 2.8 | 126 | 104 |
| S1 | ISTD1 | 99 | 93 | 89 | 93 | 2.7 | 5.6 | 4.9 | 4.7 | 4.1 | 3.1 | 0.5 | 1.5 | 115 | 109 |
| N5 | ISTD1 | 110 | 107 | 104 | 107 | 3.3 | 7.8 | 3.2 | 9.4 | 8.9 | 2.3 | 1.1 | 3.7 | 125 | 96 |
| S2 | ISTD2 | 95 | 81 | 85 | 69 | 5.1 | 8.0 | 7.4 | 11.6 | 13.7 | 10.6 | 1.4 | 4.7 | 81 | 79 |
| N7 | ISTD1 | 130 | 118 | 111 | 94 | 4.9 | 11.0 | 5.8 | 12.4 | 11.5 | 6.2 | 2.3 | 7.7 | 118 | 90 |
| O1 | ISTD1 | 104 | 92 | 99 | 88 | 2.9 | 7.5 | 4.7 | 11.7 | 9.0 | 7.6 | 2.3 | 7.6 | 85 | 83 |
| S3 | ISTD3 | 78 | 72 | 74 | 74 | 5.2 | 4.4 | 8.1 | 15.0 | 9.4 | 8.5 | 1.2 | 4.0 | 76 | 71 |
| O2 | ISTD2 | 105 | 98 | 105 | 92 | 2.7 | 4.3 | 5.1 | 8.3 | 3.3 | 2.7 | 1.0 | 3.3 | 93 | 98 |
| S4 | ISTD3 | 86 | 87 | 82 | 82 | 3.7 | 4.6 | 6.6 | 7.6 | 9.1 | 6.3 | 0.7 | 2.3 | 82 | 79 |
| N9 | ISTD3 | 95 | 89 | 96 | 96 | 2.2 | 4.0 | 4.2 | 5.3 | 7.5 | 2.8 | 0.6 | 1.8 | 105 | 110 |
| N11 | ISTD3 | 91 | 80 | 83 | 79 | 2.2 | 6.5 | 4.1 | 8.3 | 4.6 | 2.2 | 0.8 | 2.6 | 85 | 67 |
| N12 | ISTD3 | 91 | 81 | 87 | 89 | 2.7 | 3.6 | 3.0 | 4.6 | 5.4 | 0.8 | 0.4 | 1.5 | 92 | 90 |
| N13 | ISTD3 | 94 | 81 | 89 | 83 | 3.6 | 5.3 | 6.4 | 7.4 | 9.4 | 5.6 | 1.1 | 3.6 | 94 | 92 |
| S5 | ISTD3 | 79 | 85 | 75 | 69 | 3.1 | 6.2 | 10.8 | 11.8 | 11.0 | 10.2 | 0.9 | 2.9 | 85 | 81 |
| N14 | ISTD3 | 90 | 97 | 91 | 87 | 2.8 | 9.6 | 8.5 | 13.1 | 9.6 | 1.2 | 1.2 | 4.1 | 80 | 64 |
| N15 | ISTD3 | 87 | 86 | 85 | 77 | 2.7 | 10.8 | 4.2 | 8.8 | 10.3 | 5.1 | 1.2 | 3.9 | 83 | 65 |
| S6 | ISTD5 | 101 | 96 | 92 | 77 | 5.7 | 7.8 | 10.3 | 10.8 | 3.0 | 11.5 | 0.9 | 3.0 | 85 | 85 |
| N16 | ISTD5 | 100 | 72 | 73 | 78 | 3.3 | 5.0 | 7.9 | 8.3 | 8.5 | 4.8 | 0.7 | 2.4 | 87 | 84 |
| S7 | ISTD6 | 104 | 98 | 82 | 71 | 3.1 | 7.5 | 9.9 | 15.1 | 12.6 | 11.2 | 1.2 | 3.9 | 86 | 81 |
| O3 | ISTD6 | 105 | 100 | 100 | 101 | 2.0 | 16.0 | 7.0 | 9.0 | 11.9 | 10.1 | 1.0 | 3.3 | 98 | 98 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| N17 | ISTD6 | 99 | 97 | 100 | 88 | 2.1 | 15.0 | 7.5 | 7.8 | 9.2 | 8.1 | 0.7 | 2.5 | 97 | 95 |
| N18 | ISTD7 | 102 | 103 | 107 | 114 | 3.8 | 9.6 | 4.2 | 12.2 | 5.6 | 4.8 | 1.5 | 4.9 | 77 | 76 |
| S8 | ISTD6 | 100 | 100 | 95 | 84 | 2.8 | 24.9 | 8.3 | 7.6 | 9.6 | 10.4 | 0.7 | 2.3 | 102 | 99 |
| N19 | ISTD7 | 100 | 90 | 99 | 102 | 11.4 | 5.7 | 6.4 | 7.6 | 6.0 | 6.4 | 0.8 | 2.8 | 89 | 88 |
| N20 | ISTD7 | 85 | 91 | 93 | 107 | 5.5 | 7.7 | 8.9 | 13.5 | 3.4 | 3.6 | 1.5 | 5.1 | 100 | 103 |
| N21 | ISTD7 | 85 | 84 | 90 | 92 | 4.9 | 8.2 | 4.5 | 6.3 | 6.0 | 2.0 | 0.6 | 2.1 | 96 | 95 |
| S9 | ISTD9 | 93 | 94 | 101 | 108 | 1.8 | 11.7 | 8.3 | 13.2 | 11.2 | 3.4 | 1.5 | 5.1 | 81 | 81 |
| S10 | ISTD9 | 133 | 102 | 107 | 99 | 2.2 | 10.8 | 11.9 | 14.8 | 4.4 | 7.6 | 1.6 | 5.2 | 82 | 84 |
| O4 | ISTD9 | 116 | 99 | 108 | 103 | 2.8 | 5.6 | 3.3 | 5.9 | 5.7 | 4.7 | 0.6 | 2.2 | 92 | 96 |
| N22 | ISTD11 | 109 | 87 | 89 | 91 | 5.5 | 2.9 | 11 | 10.9 | 13.7 | 8.4 | 1.1 | 3.5 | 86 | 88 |
| N23 | ISTD11 | 112 | 89 | 88 | 84 | 2.4 | 5 | 12.5 | 17 | 18.2 | 24.2 | 1.5 | 5.1 | 80 | 84 |
| **Halo-HPACs** | | | | | | | | | | | | | | | |
| HH1 | ISTD2 | 67 | 87 | 82 | 77 | 4.7 | 5.7 | 7.8 | 7.7 | 5.7 | 6.8 | 0.6 | 2.1 | 56 | 54 |
| HH2 | ISTD2 | 80 | 83 | 90 | 90 | 2.7 | 7.2 | 3.5 | 6.5 | 2.4 | 3.9 | 0.6 | 2.1 | 78 | 73 |
| HH3 | ISTD2 | 91 | 90 | 99 | 97 | 2.4 | 6.3 | 2.6 | 5.6 | 3.7 | 3.8 | 0.6 | 1.9 | 89 | 87 |
| HH4 | ISTD2 | 75 | 90 | 91 | 94 | 3.1 | 4.8 | 4.3 | 7.0 | 4.4 | 2.2 | 0.7 | 2.3 | 79 | 76 |
| HH5 | ISTD2 | 77 | 85 | 90 | 88 | 3.5 | 4.5 | 3.9 | 5.2 | 1.7 | 2.2 | 0.5 | 1.6 | 78 | 75 |
| HH6 | ISTD1 | 94 | 95 | 93 | 94 | 4.0 | 7.8 | 4.9 | 7.9 | 3.4 | 2.8 | 0.9 | 3.0 | 95 | 94 |
| HH7 | ISTD2 | 79 | 67 | 61 | 59 | 4.7 | 13.0 | 10.1 | 19.5 | 2.9 | 4.7 | 1.3 | 4.2 | 77 | 78 |
| HH8 | ISTD2 | 93 | 96 | 93 | 93 | 2.5 | 7.6 | 3.5 | 6.8 | 1.7 | 3.7 | 0.8 | 2.6 | 100 | 101 |
| HH9 | ISTD1 | 94 | 84 | 67 | 67 | 6.0 | 7.3 | 5.9 | 19.4 | 2.3 | 5.2 | 1.4 | 4.7 | 83 | 85 |
| HH10 | ISTD3 | 74 | 80 | 86 | 87 | 4.1 | 7.4 | 4.6 | 6.7 | 6.4 | 6.1 | 0.6 | 2.1 | 72 | 67 |
| HH11 | ISTD3 | 43 | 35 | 48 | 63 | 19.2 | 14.1 | 9.1 | 24.6 | 17.9 | 25.3 | 1.7 | 5.6 | 91 | 80 |
| HH12 | ISTD3 | 75 | 93 | 100 | 98 | 8.2 | 10.4 | 5.7 | 11.2 | 8.3 | 1.3 | 1.2 | 3.9 | 83 | 87 |
| HH13 | ISTD3 | 42 | 37 | 47 | 43 | 20.7 | 28.0 | 12.4 | 22.8 | 11.9 | 20.5 | 1.2 | 4.0 | 74 | 74 |
| HH14 | ISTD6 | 99 | 99 | 94 | 98 | 3.1 | 8.1 | 6.6 | 10.5 | 5.6 | 5.8 | 1.1 | 3.8 | 113 | 113 |
| HH15 | ISTD6 | 99 | 87 | 97 | 88 | 3.5 | 6.8 | 4.1 | 7.2 | 6.5 | 4.7 | 0.7 | 2.3 | 101 | 102 |
| HH16 | ISTD7 | 86 | 78 | 83 | 78 | 4.6 | 9.7 | 5.1 | 9.5 | 6.3 | 7.7 | 0.8 | 2.7 | 96 | 97 |

'C1 = 800 pg/µL, C2 = 100 pg/µL, C3 = 25 pg/µL and C4 = 10 pg/µL', 'D2 = day 2 and D3 = day 3', 'P1 = dispersive mass change and P2 = ASE temperature change'.

## 3.3 RESULTS AND DISCUSSION

### 3.3.1 Mass spectral characterization

The first step in our analytical method development was to build our GC-MS/MS MRM detection and quantitation method. To do so, we first acquired full-scan EI mass spectra of all 80 standards by injecting ~1 ng/uL of each solution prepared in hexane. Because of their high boiling points, we were unable to obtain EI mass spectra for 5,6,6a,11a-tetrahydro-11H-benzo[a]carbazole, and diacenaphthyleno[1,2-b;1',2'-d]thiophene. Additionally, we omitted single ring compounds like pyrrole, pyridine, thiophene and furan and their alkylated and halogenated derivatives from our method because of their low boiling points and early chromatographic elution. Careful examination of the full-scan EI mass spectra of each compound enabled us to select a characteristic precursor ion that was purposely fragmented in the product ion scan mode to generate diagnostic product ions. Two precursor-product ion transitions formed the basis of our multiple reaction monitoring method and were used for each analyte; one to be used as a quantitative ion transition and the other as a confirmation ion transition. Because of the large number of MRM transitions and to increase instrument detectability, we used two GC-injections and two separate MRM methods. Two criteria were used to confirm the presence of our target analytes. First, the ratio of the quantitation and confirmation MRM ion transitions in the samples must be within ± 20% of their respective theoretical values. Second, the retention time ($r_t$) of our analytes must be within ±1s to those in the standard solution.

Table 3.1 shows the two MRM methods that were developed for HPACs and Halo-HPACs incorporated with respective MRM transitions of labeled and native PAHs. Figures 3.1 and 3.2 depict the gas chromatograms obtained for two of the MRM methods. 1,2,3,4-tetrahydrocarbazole

coelutes with benzo[h]quinoline and acridine at $r_t$ of 9.85 min, but because of their unique MRM ion transitions, we were able to mass resolve all 3 compounds. Similarly, while the internal standards anthracene-$d_{10}$ in the HPACs chromatogram and naphthalene-d8 in the halo-HPACs chromatogram coelute with the neighbouring peaks, we were able to mass resolve all these compounds.

### 3.3.2 Extraction optimization

Previous work in our laboratory has shown that ASE extractions using DCM result in high recoveries for some non-hetero-PACs [31]. However, the presence of the electronegative heteroatoms makes these compounds more polar relative to other polycyclic aromatic compounds that do not contain N-, O- and S. We hypothesized then that the presence of water in our samples and a less-polar solvent like DCM would both contribute to less than desired recoveries of our target analytes. Unsurprisingly, we observed low recoveries of our analytes when DCM was used as the lone extracting solvent. Adding 50% of ethyl acetate to DCM, increased in the polarity of the extraction solvent and resulted in (i) a notable improvement in the efficiency of our extraction procedure for HPACs and (ii) good recoveries for PAHs (see Table I 4). The comparison of recoveries of HPACs between DCM and 50:50 DCM: ethyl acetate mixture is given in Table I 2. Furthermore, during the method development stage, we tested extractions using the 50:50 mixture of DCM:ethyl acetate on the same sample that was wet and dry and found no differences in analyte recoveries (see Table 3.3). The wet samples were prepared by mixing 1.0 g of sediment with 1.0 ml of optima grade water. Finally, to confirm that our method was exhaustively extracting HPACs from real samples, we performed three sequential ASE extractions on the Randle Reef sediment sample (dry and wet, *n=3* in each case). Except for acridine (64%), all our other analytes were recovered greater than

70% after a single extraction (see Table 3.3). As such, our protocol was based on a single extraction of samples.

**Table 3.3:** Mean concentrations (ng/g, wet and dry weight, *n=3*) of HPACs and halo-HPACs in Randle Reef sediments and percentage of analytes observed by sequential extractions using a 50:50 mixture of DCM:ethyl acetate

| Standard Name | Dry Sediments Extraction | | | Wet Sediments Extraction | | |
|---|---|---|---|---|---|---|
| | Step 01 Conc. (ng/ g) | Step 02 Conc. (ng/ g) | Step 03 Conc. (ng/ g) | Step 01 Conc. (ng/ g) | Step 02 Conc. (ng/ g) | Step 03 Conc. (ng/ g) |
| **Quinoline** | 12 ± 6.2 (84%) | 2.2 ± 0.7 (16%) | < LOD | 12 ± 1.4 (88%) | 1.6 ± 0.5 (12%) | < LOD |
| **Isoquinoline** | 4.2 ± 2.1 (100%) | < LOD | < LOD | 9.0 ± 0.2 (90%) | 1.0 ± 0.5 (10%) | < LOD |
| **2-Methylquinoline** | 8.7 ± 4.2 (73%) | 2.1 ± 0.4 (18%) | 1.2 ± 0.1 (10%) | 20 ± 3.7 (100%) | < LOD | < LOD |
| **3-Methylisoquinoline** | 1.4 ± 0.8 (100%) | < LOD | < LOD | 1.5 ± 0.3 (100%) | < LOD | < LOD |
| **Benzo[h]quinoline** | 115 ± 15 (97%) | 2.6 ± 0.3 (2%) | 1.3 ± 0.2 (1%) | 139 ± 0.1 (99%) | 1.0 ± 0.1 (1%) | < LOD |
| **3-Methylbenzo[f]quinoline** | 18 ± 2.3 (91%) | 1.9 ± 0.4 (9%) | < LOD | 28 ± 6.5 (100%) | < LOD | < LOD |
| **2-Chloroquinoline** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **2-Bromoquinoline** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **3-Bromoisoquinoline** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **4-Phenylpyridine** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **Indole** | 2.6 ± 1.3 (100%) | < LOD | < LOD | 17 ± 4.1 (100%) | < LOD | < LOD |
| **5-Chloroindole** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **5-Bromoindole** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **Acridine** | 80 ± 5.1 (64%) | 24 ± 2.2 (19%) | 21 ± 1.7 (17%) | 311 ± 68 (98%) | 4.9 ± 0.7 (2%) | 2.4 ± 0.8 (1%) |
| **2-Methylacridine** | 56 ± 7.9 (94%) | 2.0 ± 0.2 (3%) | 1.4 ± 0.2 (2%) | 71 ± 26 (100%) | < LOD | < LOD |
| **Benz[c]acridine** | 250 ± 41 (97%) | 5.5 ± 0.2 (2%) | 2.7 ± 0.4 (1%) | 295 ± 116 (100%) | 1.1 ± 0.4 (0%) | < LOD |
| **Dibenz[c,h]acridine** | 27 ± 4.6 (100%) | < LOD | < LOD | 33 ± 11 (100%) | < LOD | < LOD |
| **9-Chloroacridine** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **9-bromoacridine** | 40 ± 3.2 (100%) | < LOD | < LOD | 71 ± 15 (100%) | < LOD | < LOD |
| **Carbazole** | 1829 ± 200 (97%) | 39 ± 2.6 (2%) | 18 ± 2.2 (1%) | 2214 ± 447 (99%) | 14 ± 1.5 (1%) | 5.8 ± 0.4 (0%) |
| **9-Methylcarbazole** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |

| | | | | | | |
|---|---|---|---|---|---|---|
| **3,6-Dimethylcarbazole** | 2.3 ± 0.6 (100%) | < LOD | < LOD | 1.6 ± 0.9 (100%) | < LOD | < LOD |
| **3-Tert-butyl-9H-carbazole** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **9-Phenylcarbazole** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **11H-Benzo[a]carbazole** | 837 ± 152 (98%) | 12.4 ± 0.9 (1%) | 6.5 ± 1.3 (1%) | 789 ± 144 (100%) | 2.9 ± 0.9 (0%) | < LOD |
| **7H-benzo[c]carbazole** | 470 ± 63 (98%) | 5.6 ± 0.2 (1%) | 2.5 ± 0.3 (1%) | **545 ± 217** (100%) | 1.1 ± 0.4 (0%) | < LOD |
| **13(H)-Dibenzo[a,i]carbazole** | 31 ± 4.8 (100%) | < LOD | < LOD | 35 ± 14 (100%) | < LOD | < LOD |
| **3,6-Dichloro-9h-carbazole** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **3,6-Dibromocarbazole** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **4-Methylbenzothiophene** | 4.3 ± 1.8 (100%) | < LOD | < LOD | 4.1 ± 0.7 (100%) | < LOD | < LOD |
| **2,3-Dimethylbenzothiophene** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **2,3,4-Trimethylbenzothiophene** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **Dibenzothiophene** | 716 ± 69 (97%) | 16 ± 1.0 (2%) | 7.6 ± 0.8 (1%) | 853 ± 163 (100%) | 2.0 ± 0.3 (0%) | < LOD |
| **4-Methyldibenzothiophene** | 33 ± 3.4 (100%) | < LOD | < LOD | 37 ± 7.0 (100%) | < LOD | < LOD |
| **2,4-Dimethyldibenzothiophene** | 3.3 ± 0.6 (100%) | < LOD | < LOD | 6.1 ± 2.5 (100%) | < LOD | < LOD |
| **2,4,7-Trimethyldibenzothiophene** | 2.8 ± 0.1 (100%) | < LOD | < LOD | 3.8 ± 0.7 (100%) | < LOD | < LOD |
| **7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **6-Methylbenzo[b]naphtho[2,1-d]thiophene** | 13 ± 2.8 (100%) | < LOD | < LOD | 15 ± 5.6 (100%) | < LOD | < LOD |
| **2,3-Dibromothiophene** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **2-Bromobenzothiophene** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **2,3-Dichlorobenzothiophene** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **2,3-Dibromobenzothiophene** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **2-Bromodibenzothiophene** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **2,8-Dibromodibenzothiophene** | < LOD | < LOD | < LOD | < LOD | < LOD | < LOD |
| **Dibenzofuran** | 1781 ± 171 (98%) | 31.0 ± 1.9 (2%) | 13.9 ± 1.4 (1%) | 1756 ± 357 (100%) | 5.9 ± 0.8 (0%) | 2.2 ± 0.6 (0%) |
| **2-Methyldibenzofuran** | 340 ± 25 (98%) | 5.7 ± 1.1 (2%) | 2.8 ± 0.4 (1%) | 336 ± 90 (100%) | 1.2 ± 0.2 | < LOD |
| **Benzo[b]naphtho[2,3-d]furan** | 326 ± 43 (98%) | 5.5 ± 0.7 (2%)) | 2.2 ± 0.3 (1%) | 356 ± 114 (100%) | < LOD | < LOD |
| **Dinaphtho[1,2-b;1',2'-d]furan** | 23 ± 5.8 (100%) | < LOD | < LOD | 27 ± 8.0 (100%) | < LOD | < LOD |

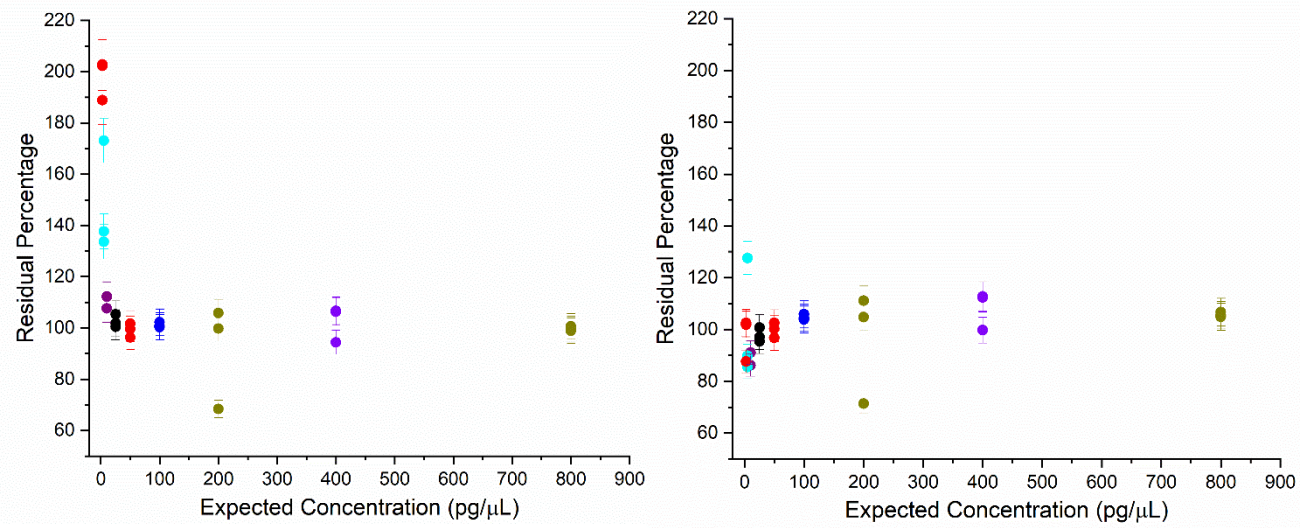



**Figure 3.3:** Residual plots of quinoline using (i) the linear regression model and (ii) the average $r_f$ model.

### 3.3.3 Performance characteristics

Although the linear regression models gave $R^2$ values greater than 0.97 for all compounds, we did observe large positive and negative bias in the residual percentage calculations for many of the compounds using this model (see Table I 5). However, we observed less bias in our data using the average $r_f$ model. Figure 3.3 shows the residual plots of quinoline using (i) the linear regression model and (ii) the average $r_f$ model. At the 2.5 pg/µL concentration level, there was a mean positive bias of 198 ± 8% (*n=3*) while the average $r_f$ gave a positive bias of 97 ± 8% (*n=3*). Similar observations were made for the other analytes. Specifically, the %RSD of the average response factors determined over the working range was less than 26% for all compounds. (see Table I 5). In light of this, the average $r_f$ was selected as our method of choice for quantitation.

Despite our best efforts, and for reasons that we are still unable to explain, three compounds (i.e., 2,5-dimethyl-1-phenylpyrrole, 2,3 dimethylindole and 1,2,3,4-tetrahydrocarbazole) consistently showed small recoveries and were purposely omitted from our validation study. Except for indole,

5-chloroindole, 9-chloroacridine and 9-bromoacridine, all our analytes had recoveries greater than 67% at all four fortification levels (10, 25, 100, and 800 pg/µL). Thirty-two standards have recoveries between 80% to 115%.

The accuracy, reported as a percent recovery, at each concentration level (10, 25, 100, and 800 pg/µL) is presented in Table 3.2. The repeatability of our method, expressed as the percent relative standard deviation (%RSD) of eight replicates, was smaller than 21% for all our target analytes with the exception of 2-methylquinoline (22%), 2-methylacridine (24%) and 13(H)-Dibenzo[a,i]carbazole (23%). The %RSD calculated to examine the intermediate precision was under 15% for analytes other than 13(H)-dibenzo[a,i]carbazole, 9-chloroacridine and 9-bromoacridine. However, only twenty-eight standards have an acceptable intermediate precision based on the results of single-factor ANOVA tests ($\alpha = 0.05$). When all the standards were considered, the LODs and LOQs ranged from 0.4 to 2.3 ng/g and from 1.5 to 7.7 ng/g, respectively. Forty-six standards showed recoveries between 65% to 125% for the extractions done by changing the two parameters in the ruggedness check (*see* Table 3.2). Mean concentrations and % RSD of analytes in the six procedural blank samples are mentioned in Table I 6.

A sediment sample collected from Randle Reef was analysed using the validated method to determine concentrations of HPACs and compare them to the PAHs and APAHs. Measurements on PAHs and APAHs (*see* Table I 7) were performed using the method described by Idowu et *al*. [32]. The sample extraction and purification procedure used for PAHs and APAHs is mentioned by Xia et al. [31]. Blank corrected concentrations (µg/g, dry weight) of analytes detected are shown in Figure 3.4. The ASE cells for procedural blanks were prepared as any other sample, without including the sediment samples. Then, those cells were spiked with internal standards.

Concentrations of analytes detected in our procedural blanks are given in Table I 8. Of the 48 HPACs in our method, 27 were detected in the sediment sample. Unsurprisingly, the concentrations of the 16 parent PAHs ($\Sigma_{16}$ = 80.9 ± 8.7 µg/g, dry weight, *n=3*) were significantly greater ($p<0.05$) than all the other PACs and account for *ca*. 80% of all PACs measured in the sample. The rank order of the other PACs measured was: $\Sigma_{30}$APAH > $\Sigma_{17}$PANH >$\Sigma_{4}$PAOH >$\Sigma_{6}$PASH. Taken together, the $\Sigma_{27}$HPACs (mean: 7.3 µg/g, dry weight) account for *ca*. 7% of all PACs measured in the sample.

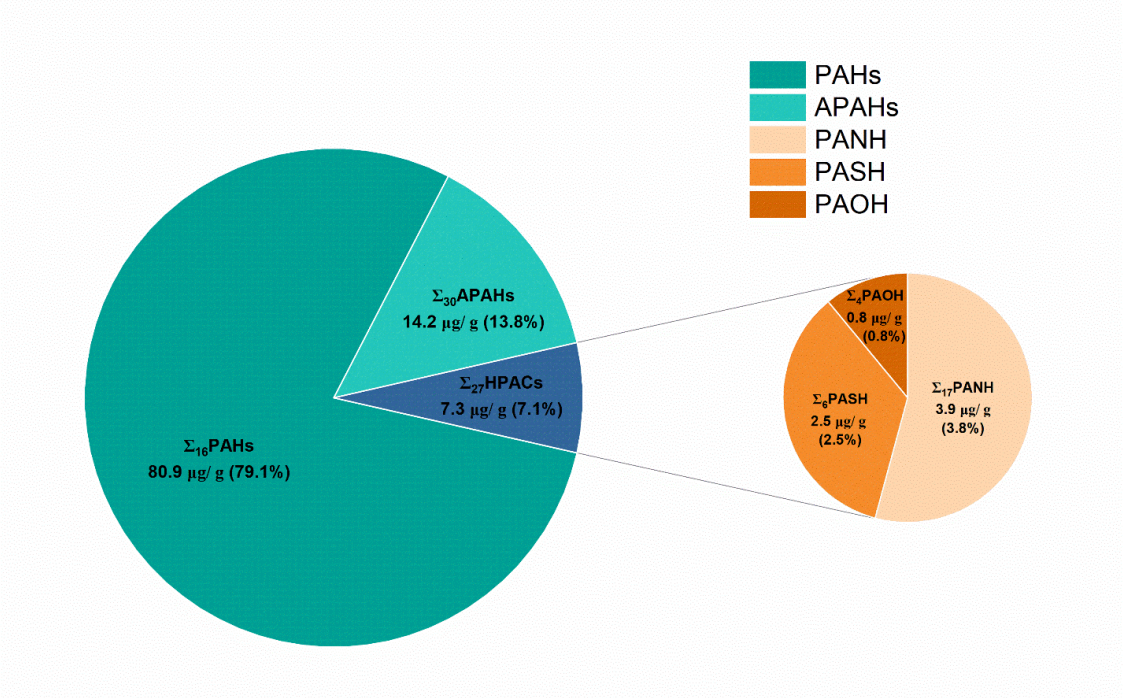



**Figure 3.4:** Blank corrected concentrations (µg/g, dry weight, *n=3*) of PACs and their relative distributions in sediments from Randal Reef (Hamilton Harbour, ON, Canada).

## 3.4 CONCLUSIONS

To date, the simultaneous quantitative analysis of nitrogen, oxygen, and sulfur-containing hetero-polycyclic aromatic compounds has been limited to only a few compounds. The targeted GC-MS/MS methods presented here encompass forty-eight non- and halogenated-based HPACs. The developed extraction and purification technique proved to be efficient and effective for the analysis of HPACs in sediment samples and was not affected by the presence of water in the samples. The empirically derived data quality objectives demonstrate that the method is fit for its intended purpose. The method was used to provide evidence of these compounds in sediments from a contaminated site in Canada. The environmental fate, distribution and the toxicological significance are some of the research questions that need to be answered in future work related to these compounds.

## 3.5 ACKNOWLEDGEMENTS

This work was supported by a Natural Sciences and Engineering Research Council Discovery Grant to GTT (Grant number: RGPIN-05354-2019). Harold Malle (Environment and Climate Change Canada) and Ken Middlebrook (Proficiency Testing Canada) providing information on logistics, collection and handling of the sediment sample from Randal Reef (Hamilton Harbour, ON, Canada).

## 3.6 REFERENCES


(1) Abdel-Shafy, H. I.; Mansour, M. S. M. A Review on Polycyclic Aromatic Hydrocarbons: Source, Environmental Impact, Effect on Human Health and Remediation. *Egyptian Journal of Petroleum.* Egyptian Petroleum Research Institute March 1, **2016**, pp 107–123. https://doi.org/10.1016/j.ejpe.2015.03.011.

(2) Marvin, C. H.; Berthiaume, A.; Burniston, D. A.; Chibwe, L.; Dove, A.; Evans, M.; Hewitt, L. M.; Hodson, P. V.; Muir, D. C. G.; Parrott, J.; Thomas, P. J.; Tomy, G. T. Polycyclic Aromatic Compounds in the Canadian Environment: Aquatic and Terrestrial Environments. *Environmental Pollution.* Elsevier Ltd September 15, **2021**. https://doi.org/10.1016/j.envpol.2021.117442.

(3) Heravi, M. M.; Zadsirjan, V. Prescribed Drugs Containing Nitrogen Heterocycles: An Overview. *RSC Advances*. Royal Society of Chemistry November 23, **2020**, pp 44247–44311. https://doi.org/10.1039/d0ra09198g.

(4) Lellis, B.; Fávaro-Polonio, C. Z.; Pamphile, J. A.; Polonio, J. C. Effects of Textile Dyes on Health and the Environment and Bioremediation Potential of Living Organisms. *Biotechnology Research and Innovation* **2019**, 3 (2), 275–290. https://doi.org/10.1016/j.biori.2019.09.001.

(5) Kimbrough, K.; Jacob, A.; Regan, S.; Davenport, E.; Edwards, M.; Leight, A. K.; Freitag, A.; Rider, M.; Johnson, W. E. Characterization of Polycyclic Aromatic Hydrocarbons in the Great Lakes Basin Using Dreissenid Mussels. *Environ. Monit. Assess*. **2021**, 193 (12). https://doi.org/10.1007/s10661-021-09401-7.

(6) Wnorowski, A.; Aklilu, Y. abeba; Harner, T.; Schuster, J.; Charland, J. P. Polycyclic Aromatic Compounds in Ambient Air in the Surface Minable Area of Athabasca Oil Sands in Alberta (Canada). *Atmos. Environ*. **2021**, 244. https://doi.org/10.1016/j.atmosenv.2020.117897.

(7) Yu, Y.; Katsoyiannis, A.; Bohlin-Nizzetto, P.; Brorström-Lundén, E.; Ma, J.; Zhao, Y.; Wu, Z.; Tych, W.; Mindham, D.; Sverko, E.; Barresi, E.; Dryfhout-Clark, H.; Fellin, P.; Hung, H. Polycyclic Aromatic Hydrocarbons Not Declining in Arctic Air Despite Global Emission Reduction. *Environ. Sci. Technol*. **2019**, 53 (5), 2375–2382. https://doi.org/10.1021/acs.est.8b05353.

(8) Hites, R. A. Polycyclic Aromatic Hydrocarbons in the Atmosphere near the Great Lakes: Why Do Their Concentrations Vary? *Environ. Sci. Technol*. **2021**, 55 (14), 9444–9449. https://doi.org/10.1021/acs.est.0c06905.

(9) Xia, Z.; Idowu, I.; Marvin, C.; Thomas, P. J.; Johnson, W.; Francisco, O.; Stetefeld, J.; Crimmins, B.; Fry, M.; Tomy, G. T. Identification of Halogenated Polycyclic Aromatic Hydrocarbons in Biological Samples from Alberta Oil-Sands Region. *Chemosphere* **2019**, 215, 206–213. https://doi.org/10.1016/j.chemosphere.2018.10.050.

(10) Toxic substances list: schedule 1. Canadian Environmental Protection Act 1999, https://www.canada.ca/en/environment-climate-change/services/canadian-environmental-protection-act-registry/substances-list/toxic/schedule-1.html.

(11) Filewood, T.; Kwok, H.; Brunswick, P.; Yan, J.; Ollinik, J. E.; Cote, C.; Kim, M.; van Aggelen, G.; Helbing, C. C.; Shang, D. Investigating the Fate of Polycyclic Aromatic Sulfur Heterocycle Compounds in Spilled Oils with a Microcosm Weathering Experiment. *Environmental Systems Research* **2022**, 11 (1). https://doi.org/10.1186/s40068-022-00252-w.

(12) Ghosh, P.; Mukherji, S. Fate, Detection Technologies and Toxicity of Heterocyclic PAHs in the Aquatic and Soil Environments. *Science of the Total Environment*. Elsevier B.V. September 20, **2023**. https://doi.org/10.1016/j.scitotenv.2023.164499.

(13) Andersson, J. T.; Achten, C. Time to Say Goodbye to the 16 EPA PAHs? Toward an Up-to-Date Use of PACs for Environmental Purposes. *Polycycl. Aromat. Compd.* **2015**, 35 (2–4), 330–354. https://doi.org/10.1080/10406638.2014.991042.

(14) Peddinghaus, S.; Brinkmann, M.; Bluhm, K.; Sagner, A.; Hinger, G.; Braunbeck, T.; Eisenträger, A.; Tiehm, A.; Hollert, H.; Keiter, S. H. Quantitative Assessment of the Embryotoxic Potential of NSO-Heterocyclic Compounds Using Zebrafish (Danio Rerio). *Reproductive Toxicology* **2012**, 33 (2), 224–232. https://doi.org/10.1016/j.reprotox.2011.12.005.

(15) Comber, M. H. I.; De Ferrer, J. A.; Djemel, N.; Eadsforth, C. V; Lescrauwaet, A.; Léon, M.; Linington, P. S.; Redman, A.; Villalobos, S. A.; Den Haan, K. H. Analysis of N-, O-, and S-Heterocyclics in Petroleum Products Using GCxGC with Specific Detection; **2016**. www.concawe.org.

(16) Dias, L. C.; Garcia, K. S.; de Souza Queiroz, A. F.; Machado, M. E. Quantification of Nitrogen Markers in Crude Oils Using Gas Chromatography-Triple Quadrupole Mass Spectrometry. *Fuel* **2021**, 305. https://doi.org/10.1016/j.fuel.2021.121477.

(17) Ohura, T.; Kitazawa, A.; Amagai, T.; Makino, M. Occurrence, Profiles, and Photostabilities of Chlorinated Polycyclic Aromatic Hydrocarbons Associated with Particulates in Urban Air. *Environ. Sci. Technol.* **2005**, 39 (1), 85–91. https://doi.org/10.1021/es040433s.

(18) Ni, H. G.; Zeng, E. Y. Environmental and Human Exposure to Soil Chlorinated and Brominated Polycyclic Aromatic Hydrocarbons in an Urbanized Region. *Environ. Toxicol. Chem.* **2012**, 31 (7), 1494–1500. https://doi.org/10.1002/etc.1854.

(19) Sun, J. L.; Ni, H. G.; Zeng, H. Occurrence of Chlorinated and Brominated Polycyclic Aromatic Hydrocarbons in Surface Sediments in Shenzhen, South China and Its Relationship to Urbanization. *Journal of Environmental Monitoring* **2011**, 13 (10), 2775–2781. https://doi.org/10.1039/c1em10465a.

(20) Shang, D.; Kim, M.; Haberl, M. Rapid and Sensitive Method for the Determination of Polycyclic Aromatic Hydrocarbons in Soils Using Pseudo Multiple Reaction Monitoring Gas

Chromatography/Tandem Mass Spectrometry. *J. Chromatogr. A* **2014**, 1334, 118–125. https://doi.org/10.1016/j.chroma.2014.01.074.

(21) Machado, M. E.; Nascimento, M. M.; Bomfim Bahia, P. V.; Martinez, S. T.; Bittencourt de Andrade, J. Analytical Advances and Challenges for the Determination of Heterocyclic Aromatic Compounds (NSO-HET) in Sediment: A Review. *TrAC - Trends in Analytical Chemistry.* Elsevier B.V. May 1, **2022**. https://doi.org/10.1016/j.trac.2022.116586.

(22) Droppo, I. G.; di Cenzo, P.; Power, J.; Jaskot, C.; Chambers, P. A.; Alexander, A. C.; Kirk, J.; Muir, D. Temporal and Spatial Trends in Riverine Suspended Sediment and Associated Polycyclic Aromatic Compounds (PAC) within the Athabasca Oil Sands Region. *Science of the Total Environment* **2018**, 626, 1382–1393. https://doi.org/10.1016/j.scitotenv.2018.01.105.

(23) Manzano, C. A.; Marvin, C.; Muir, D.; Harner, T.; Martin, J.; Zhang, Y. Heterocyclic Aromatics in Petroleum Coke, Snow, Lake Sediments, and Air Samples from the Athabasca Oil Sands Region. *Environ. Sci. Technol.* **2017**, 51 (10), 5445–5453. https://doi.org/10.1021/acs.est.7b01345.

(24) Chibwe, L.; Roberts, S.; Shang, D.; Yang, F.; Manzano, C. A.; Wang, X.; Kirk, J. L.; Muir, D. C. G. A One-Century Sedimentary Record of N- and S-Polycyclic Aromatic Compounds in the Athabasca Oil Sands Region in Canada. *Chemosphere* **2020**, 260. https://doi.org/10.1016/j.chemosphere.2020.127641.

(25) Berthiaume, A.; Galarneau, E.; Marson, G. Polycyclic Aromatic Compounds (PACs) in the Canadian Environment: Sources and Emissions. *Environmental Pollution*. Elsevier Ltd January 15, **2021**. https://doi.org/10.1016/j.envpol.2020.116008.

(26) Hodson, P. V.; Wallace, S. J.; de Solla, S. R.; Head, S. J.; Hepditch, S. L. J.; Parrott, J. L.; Thomas, P. J.; Berthiaume, A.; Langlois, V. S. Polycyclic Aromatic Compounds (PACs) in the Canadian Environment: The Challenges of Ecological Risk Assessments. *Environmental Pollution.* Elsevier Ltd November 1, **2020**. https://doi.org/10.1016/j.envpol.2020.115165.

(27) Yang, C.; Lambert, P.; Nguyen, M.; Yang, Z.; Hollebone, B. P.; Fieldhouse, B.; Brown, C. E. Application of Gas Chromatography-High Resolution Quadrupole Time-of-Flight Mass Spectrometry in Fingerprinting Analysis of Polycyclic Aromatic Sulfur Heterocycles. *J. Chromatogr. A* **2020**, 1630. https://doi.org/10.1016/j.chroma.2020.461577.

(28) Li, L.; Jiang, L.; George, S. C.; Liu, Z. Aromatic Compounds in Lacustrine Sediments from the Lower Cretaceous Jiufotang Formation, Chaoyang Basin (NE China). *Mar. Pet. Geol.* **2021**, 129. https://doi.org/10.1016/j.marpetgeo.2021.105111.

(29) Bahia, P. V. B.; Nascimento, M. M.; Hatje, V.; de Andrade, J. B.; Machado, M. E. Microscale Extraction Combined with Gas Chromatography/Mass Spectrometry for the Simultaneous Determination of Polycyclic Aromatic Hydrocarbons and Polycyclic Aromatic Sulfur Heterocycles in Marine Sediments. *J. Chromatogr. A* **2021**, 1653. https://doi.org/10.1016/j.chroma.2021.462414.

(30) Magnusson, B., Örnemark, U., Eds. The Fitness for Purpose of Analytical Methods : A Laboratory Guide to Method Validation and Related Topics, Second edition, **2014**.

(31) Xia, Z.; Idowu, I.; Kerr, E.; Klaassen, N.; Assi, H.; Bray, H.; Marvin, C.; Thomas, P. J.; Stetefeld, J.; Tomy, G. T. New Approaches to Reduce Sample Processing Times for the Determination of Polycyclic Aromatic Compounds in Environmental Samples. *Chemosphere* **2021**, 274. https://doi.org/10.1016/j.chemosphere.2021.129738.

(32) Idowu, I.; Francisco, O.; Thomas, P. J.; Johnson, W.; Marvin, C.; Stetefeld, J.; Tomy, G. T. Validation of a Simultaneous Method for Determining Polycyclic Aromatic Compounds and Alkylated Isomers in Biota. *Rapid Communications in Mass Spectrometry* **2018**, 32 (3), 277–287. https://doi.org/10.1002/rcm.8035.

# CHAPTER 4 HETEROCYCLIC AROMATIC COMPOUNDS IN LAKE ONTARIO SEDIMENTS: SPATIOTEMPORAL DISTRIBUTION AND NEW APPROACHES TO FINGERPRINTING

This chapter presents my approach to identify sources of PACs in lake sediments using advanced chemometric techniques such as PCA and UMAP. I first established chemical profiles of PACs in suspended and bottom sediments from three regions of Lake Ontario - Hamilton Harbour (HH), Toronto Harbour (TH), and non-harbour areas (NH) and compared them to commercially available coal-tar and non-coal-tar products which are known to be sources of polycyclic aromatic hydrocarbons in this aquatic system. In the absence of empirical chemical persistence measurements, I devised an ecological tool to assess chemical persistence of HPACs which was instrumental in helping to discern sources and spatial distribution of PACs in this region.

A version of this chapter has already been published as Vitharana Nipuni N, Halldorson Thor, Dereviankin Mike, Paddock Brett, Gerring Aviva, Marvin Chris, Helm Paul, Thomas Philippe J, Idowu Ifeoluwa Grace, Tomy Gregg T, Heterocyclic aromatic compounds in Lake Ontario sediments: Spatiotemporal distribution and new approaches to fingerprinting, Science of The Total Environment, 2025, 986, 179802. Copyright © 2025. B.V. Elsevier. Reprinted with permission.

## 4.1 ABSTRACT

Hetero-polycyclic aromatic compounds (HPACs), an understudied subclass of polycyclic aromatic compounds (PACs), are characterized by the replacement of one or more carbon in their aromatic ring structure with nitrogen, sulfur, and/or oxygen. This study examined the spatiotemporal distribution of fifty HPACs in sediment samples from Hamilton Harbour (HH), Toronto Harbour (TH), and nearby non-Harbour areas of Lake Ontario, proposing new approaches to fingerprinting PACs using HPACs. Although no temporal trends in HPAC concentrations were observed from 2014 to 2023, notable spatial differences were evident, with the rank order of total HPAC concentrations as HH > TH > non-harbour areas. Based on their higher source specificity relative to polycyclic aromatic hydrocarbons (PAHs), chemical fingerprinting using diagnostic ratio analysis, compositional analysis, and multivariate analysis was performed. Results from an empirical approach, environmental data, and physical-chemical properties were used to select analytes for fingerprinting. Diagnostic ratios between alkylated dibenzothiophenes and parent dibenzothiophene indicate lower ratios for coal tar contaminated sites than other sites, suggesting pyrogenic sources. Spearman's correlation between total combustion PAHs and HPACs indicates most analytes were derived from combustion processes. Principal Component Analysis (PCA) did not differentiate between sample HPAC fingerprints; samples separated in the first principal component were driven by total HPAC concentration and not variability in source fingerprint. Uniform Manifold Approximation and Projection (UMAP) provided a more detailed visualization, highlighting compositional differences between harbour sediments and non-harbour sites. Overall, these results highlight the usefulness and importance of using combined analytical approaches to fingerprint PACs using HPACs as chemical markers.

## 4.2 INTRODUCTION

Due to their persistence, bioaccumulation and toxicity, polycyclic aromatic compounds (PACs) are of environmental and health concern and understanding their sources, distribution, and behaviour is important for mitigating their impacts on ecosystems and human health.[1,2] Both the Canadian Environmental Protection Act, 1999 and the United States Environmental Protection Agency (US EPA) list 16 polycyclic aromatic hydrocarbons (PAHs) as priority pollutants.[3] However, because of the inherent chemical complexity of PACs, it has been argued that understanding the environmental fate and behavior of other classes of PACs is needed.[4,5]

Hetero-polycyclic aromatic compounds (HPACs) are a subclass of PACs that have at least one nitrogen, sulfur and/or oxygen atom replacing one or more carbon atoms in the aromatic structure. Despite sharing similar origins, simultaneous analysis of nitrogen, sulfur, and oxygen-based parent HPACs and their derivatives has received considerably less scientific attention than PAHs and other PACs, which is due, in part, to the large number of theoretical structural isomers and derivatives possible.[6,7]

Polycyclic aromatic compounds are produced either due to anthropogenic activities such as industrial and automobile emissions, tobacco smoke, agricultural discharges, or natural causes such as forest fires and microbial degradation of larger organic molecules.[7–9] It is well established that PACs enter the environment from pyrogenic, petrogenic and biogenic sources[10] and teasing out contributions from each of these origins is the focus of source identification studies. For example, some PAHs and dibenzothiophenes have been used as chemical markers for fingerprinting of PACs to the environment.[11,12] However, there is a newfound interest in the use of HPACs for environmental forensic studies due to their higher source specificity relative to

native PAHs. For example, benzocarbazole isomers have been used as a selective chemical fingerprint for samples contaminated by coal-tar-based products.[13,14]

Several factors need to be considered to ensure reliable source attribution by chemical fingerprinting. Some of the more notable factors include accurate analytical methods to characterize source materials and samples from the receiving environment, mathematical tools for data analysis, and a careful selection of chemicals used for fingerprinting.[15,16] Chemicals that are environmentally persistent and resist chemical and/or biological transformation from source to receptor are paramount for accurate fingerprinting studies. In the absence of experimentally derived persistence data, predictive environmental fate and behaviour models are usually employed to assign chemical persistence.

Sediments serve as the primary repository for most persistent organic pollutants (POPs) that enter aquatic systems. Monitoring programs often use sediments to assess regional and intercontinental contamination of POPs and to determine sediment data quality objectives. In Lake Ontario, studies have focused on analysing one or a few types of common HPACs such as dibenzothiophenes, carbazoles, and dibenzofurans, while others have examined halogenated HPACs like halogenated carbazoles.[17,18] However, expanding the suite of analytes even further will allow us to: (i) better characterize the extent of HPAC contamination in Lake Ontario and (ii) improve our understanding of the source of inputs into the system.

Additional PACs, including polyaromatic nitrogen heterocycles (PANHs), polyaromatic sulfur heterocycles (PASHs), and polycyclic oxygen heterocycles (PAOHs) and their halogenated derivatives, are presumably present in Great Lake sediments. The spatio-temporal distribution of these PACs was assessed in sediment samples collected from three geographic regions in Lake

Ontario: Hamilton Harbour (HH), Toronto Harbour (TH) and various non-harbour areas. Concentrations of HPACs in 14 different reference materials were also analyzed and were grouped into three categories: coal tar-based, non-coal tar-based, and mixed. The fingerprinting analysis was performed using diagnostic ratio analysis, compositional analysis, and multivariance techniques such as principal component analysis (PCA) and uniform manifold approximation and projection (UMAP).

## 4.3 MATERIALS AND METHOD

### 4.3.1 Chemicals

All polycyclic aromatic analytes were purchased from Chiron (AS Stiklestadveien 1, 7041 Trondheim, Norway) and Sigma Aldrich Canada Ltd (Oakville, Ontario, Canada). Thirty-one N-based, fourteen S-based, and four O-based HPACs were used in our study. High-purity (Optima grade) organic solvents, silica gel (923 grade, 100 - 200 mesh), alumina (60 - 325 mesh), Ottawa sand and anhydrous sodium sulfate, diatomaceous earth (DE) were purchased from Fisher Chemicals (Ottawa, Canada). Deuterated PAHs used as internal standards were purchased from Chiron. All the target analytes and internal standards used in this study are summarized in Table II 1.

### 4.3.2 Reference materials

Fourteen commercially available reference materials were used to compare their HPAC profiles with the sediment samples analyzed in this study. The reference materials included tire rubber which are *ca*. 20 years and 3 years old, polymer coal tar seal coat, asphalt-based sealer, airport seal coat, coal tar emulsion sealer, a chunk of coal tar from a transport truck spill that occurred in 2015, tar from tar and chip roadway collected from 2015, railway ties pieces, asphalt shingle pieces,

airport grade driveway re-surfacer-repair, non-fibrous roof coating, roof tar used on schools and crack filler -liquid. The reference materials were extracted with a procedure that ensures the optimum extraction of targeted HPACs.

### 4.3.3 Sediment samples

Sixty-nine sediment samples from 24 different sampling sites in Lake Ontario were collected between 2014 and 2023. Sampling sites were separated into three groups: Hamilton Harbour (HH), Toronto Harbour (TH) and sites upstream of river mouths, located outside of the two harbours (non-harbour sediments - NH). The HH1 to HH7 are suspended sediments, while the remaining samples (HH8 – HH10, all the TH samples, and NH) are bottom sediments. The bottom sediments are from 2014 and/or 2018, and the suspended sediments were from 6 different time points (2014, 2017, 2018, 2019, 2022 and 2023). Table II 2 and Figure II 1 depict the sample names, sampling year, and their geographical locations. The suspended sediments were collected from seven different stations located in Hamilton harbour using sediment traps for over a month. These samples are from a long-term program that has been collecting Lake Ontario sediments seasonally over the period of May–November at stations throughout the harbour since 1987. More information on sample collection and preparation can be found in Burniston et al.[19] These sites reflect a variety of conditions dictated by water exchange with Lake Ontario and point- and non-point source discharges. Information about sample collection and sampling sites for bottom sediments can be found in Zhang et al.[14]

### 4.3.4 Extraction and cleanup procedures

Sample extraction and clean-up were done using the method described by Vitharana *et al*.[20] After spiking with recovery internal standard (RIS, 10 ng/uL, 10 uL), samples were extracted using

accelerated solvent extraction (ASE) with a mixture of 1:1 DCM: ethyl acetate as the extracting solvent. Extracts were then reduced to 1 mL and spiked with instrument performance internal standard (IPIS, $d_8$-anthracene, 10 ng/μL, 10 μL) before being injected into gas chromatography tandem mass spectrometry (GC/MS-MS). Reference materials were extracted using a custom made 50 mL stainless steel extraction cell made for a Precellys Evolution Homogenizer (Bertin Technologies, Montigny-le-Bretonneux, France). 0.1 g of each of the 14 coal tar reference materials was added to separate cells, spiked with RIS (10 ng/μL, 10 μL), and extracted separately with 24 mL of 1:1 DCM: ethyl acetate for 5 minutes at 1000 g. Extracts were cleaned twice using 20 g of silica gel in a dispersive solid-phase extraction setup. Silica gel was pre-washed with 20 mL of DCM and then 20 mL of ethyl acetate, air-dried and then added to a 75 mL polypropylene cartridge with DCM. Samples were added to the head of the cartridge, and 30 mL of DCM was added and collected (Fraction 1), followed by the addition of 20 mL ethyl acetate (Fraction 2). Extracts were combined and reduced in volume and further purified using gel permeation chromatography using a mixture of 50:50 DCM and hexane. The GPC extracts were reduced to 2 mL in 250 mL flat bottom flasks using 70:30 Hexane: DCM and a mixture of 4 g of silica gel, 0.5 g sodium sulfate and 1g of 5% deactivated alumina was added to each flask.[21] Flasks were allowed to sit for 30 min with intermittent swirling. Then the extracts were transferred to 60 mL ASE tubes using one 5 mL 70:30 Hexanes: DCM and two 10 mL ethyl acetate washes. Volumes were further reduced to 1 mL prior to spiking with IPIS (10 ng/uL, 10 uL) and GC/MS-MS analysis.

### 4.3.5 Gas chromatography mass spectrometry

Quantitative targeted analysis was carried out at low m/z resolution using an Agilent 7890 GC coupled with a triple quadrupole Agilent 7000C mass spectrometer (GC/MS-MS) with an electron

ionization (EI) source. The instrument parameters and details of the multiple reaction monitoring methods (MRM) are described by Vitharana *et al*.[20] High m/z resolution was used for analyte confirmation and performed using an Agilent 7890A GC coupled with a Pegasus GC-HRT 4D (LECO, St Joseph, MI, USA) high-resolution time-of-flight mass spectrometer (TOF-MS). Injections of 2 µL were carried out at 300 °C in splitless mode. A Rxi -5Sil MS (60 m × 0.25 mm × 0.25 m) (Restek, Bellefonte, PA, USA) column was used with helium as the carrier gas. The oven was programmed from 80 ˚C (held for 1 min), heated to 210 ◦C at 35 ˚C/ min, further to 260 ˚C at 3 ˚C/ min and finally to 315 ˚C at 10 ˚C/ min (held for 5 min). The mass spectrometer was run at a resolving power (full-width half height) of 25,000 based on the peak width of m/z 218.9851 of perfluorotributylamine (PFTBA). During the chromatographic run, PFTBA was continuously bled into the ion source to serve as an internal standard for instrument mass calibration. Mass accuracies of ± 5 ppm between accurately measured m/z values and theoretical m/z values were used for positive confirmation of our analytes.

### 4.3.6 Quality control

Details of the validated method used in this study can be found in Vitharana *et al*.[20] For every 20 sediment samples, one procedural blank, one spiked control sample, and one non-spiked control sample were processed. Two procedural blanks were processed with reference materials. Between every 25 GC-MS/MS injections, solvent blanks were run. The RIS recoveries of each sample were calculated using IPIS to calculate extraction efficiencies. The average recovery of the RIS was 98 ± 19 % (n=82). Negligible amounts of the analytes were detected in the procedural blanks.

### 4.3.7 Fingerprinting analysis using principal component analysis and uniform manifold approximation and projection

Statistical analysis was performed using RStudio (2024.09.0+375 "Cranberry Hibiscus" Release (c8fc7aee6dc218d5687553f9041c6b1e5ea268ff, 2024-09-16) for Windows) software. After selecting the suitable analytes for the fingerprinting studies based on their chemical persistence in sediments, the concentration values less than the lower detection limit (LOD) were replaced by half the detection limit. A binary table was constructed for each analyte for all samples to identify the detects and non-detects. For a particular site, the percent detection of each analyte was determined, and the overall detection percentage was calculated. The analytes with 90% or more non-detection percentages were excluded. The detection percentages for each analyte can be found in Table II 4. Instead of concentrations, percentage contributions of each analyte to the total concentration were used for multivariant analysis. Twenty-two analytes were used for the PCA and UMAP analysis. Prior to the PAC analysis, in order to enhance the comparability and minimize the sample-to-sample variance, data were percent normalized by dividing each analyte concentration by the total selected HPACs concentration of that specific site and then taking the percentages. The hyperparameters used for UMAP analysis were n_neighbors = 20, min_dist = 0.3, and n_components = 3.[22]

## 4.4 RESULTS AND DISCUSSION

### 4.4.1 Spatial distribution of HPACs

Twenty-two non-halogenated and nine halogenated nitrogen-containing HPACs were analysed in each sample (see Figure 4.1 and Table II 4). For HH suspended sediments collected in 2014, the concentration of $\Sigma_{31}$PANHs were highest from Windemere while the lowest $\Sigma_{31}$PANHs concentrations were measured in sediments from western part of the HH near Cootes (station 9031). These variations can be attributed to the influence of different anthropogenic activities, circulation patterns in the harbour, and proximity to sources. Windemere 914 site is in an area with high urban runoff and waste water treatment plant (WWTP) discharge outfall whereas the HH site near West Cootes is situated closer to areas less impacted by human activities and major circulation patterns in the harbour. [23] Sample locations situated near Randle Reef (HH 9083) also showed relatively high concentrations of $\Sigma_{31}$PANHs. In the most recent years, concentrations of $\Sigma_{31}$PANHs were highest in sediments from Windemere, while relatively similar $\Sigma_{31}$PANHs concentrations were observed in sediments from the other sites. The concentrations of $\Sigma_{31}$PANHs in sediments from TH were similar for all three sites. Compared to HH suspended sediments the bottom sediment samples had *ca*. 5× lower concentrations of $\Sigma_{31}$PANHs.

Indole was detected at concentrations *ca*. 10× greater than any other analyte in HH suspended sediments and similar in concentrations to PAHs and APAHs. Indole is a two-ring PANH that has a pyrrole ring attached to a benzene ring. Indole enters the environment from sources such as coal tar-based industries, tobacco smoke and animal and human waste.[24] Elevated concentrations of indole were consistent in all suspended sediments and positive detection was confirmed with HRMS. The mass accuracy calculated for indole and other analytes using high resolution TOFMS

can be found in Table II 5. Interestingly, concentrations of the indole derivatives 2,3-dimethyl indole, 5-chloroindole and 5-bromoindole were similar in concentration to other HPACs. Compared to suspended sediments bottom sediment samples from HH had much smaller concentrations of indole (*ca*. 100-1000×) which is consistent with the low environmental persistence of this compound.[25] These levels were consistent with TH bottom sediment samples.

Other two-ring PANHs, such as quinoline, isoquinoline, and their alkylated and halogenated derivatives, were detected in most of the samples (range: 82.0 to 0.4 ng/ g). Of the three-ring PANH, carbazole had the second greatest concentration in HH suspended sediments (116 ± 60 ng/g). Acridine and benzo(h)carbazole like other three ring systems and their derivatives were also present in detectable quantities in most samples (97 to 4 ng/g). Some of the four-ring and five-ring nitrogen containing compounds were also detected at high concentrations (217.0 to 0.8 ng/g). Compared to non-halogenated PANHs, the halogenated PANHs were below detection limits in many samples except for 9-bromoacridine and 3,6-dichloro-9h-carbazole.

In 2018, $\Sigma_{31}$PANHs concentrations were highest for sediments from the Oakville Index station and lowest for Rough and Don Rivers. The data indicate that indole concentrations are much smaller (*ca*. 100 to 1000×) than what was observed in HH suspended sediments. In addition to indole, the two ring systems quinoline and isoquinoline and unsubstituted three ring PANHs like carbazole and benzo(h)quinoline were detected at smaller concentrations in areas outside the two harbours confirming the continued input of these low persistent compounds. The halo-HPACs, 9-chloroacrdine, 9-bromoacridine and 3,6-dichloro-9h-carbazole were the only halogenated analytes detected in these samples and their concentration ranged from 90 to 2 ng/ g. These results agree well with halo-HPACs results from HH sediments.

As depicted in Figure 4.2, the concentrations of 15 PASHs and 4 PAOHs were detected in all samples. Concentrations of $\Sigma_{15}$PASHs were highest in sediments collected from HH at site 9033. This was consistent for both suspended sediments and bottom sediments. In more recent years, $\Sigma_{15}$PASH concentrations in suspended sediments declined slightly at all sites. Site 9031 located in a less impacted area compared to sites on the adjacent side of the harbour, had the lowest $\Sigma_{15}$PASH concentrations among all HH samples and across all years. Concentrations were relatively similar at sites in HH and TH. In non-harbour samples, concentrations of $\Sigma_{15}$PASHs were *ca*. 10× smaller than the harbour samples. Compared to $\Sigma_{31}$PANHs, $\Sigma_{15}$PASH concentrations were smaller at all the geographical locations.

Compared to benzothiophenes, dibenzothiophene derivatives were found at greater concentrations. This trend is true for samples from both geographical areas. Based on our findings, 2,3,4-trimethylbenzothiophene and 7,8,9,10-tetrahydrobenzo[b]naphtho-[2,3-d]thiophene were below detection limits in most samples. Compared to halogenated PANHs, halogenated PASH concentrations were at non-detectable levels in all samples.

The highest $\Sigma_4$PAOH concentrations were measured in Windemere samples and Randle Reef samples while the lowest concentrations were detected in HH suspended sediments at site 9031. All four oxygen-containing HPACs were detected in samples from all geographical areas. The concentrations of $\Sigma$ $\Sigma_4$PAOHs were similar in all samples and ranged from 280 to  6 ng/ g. Similar to $\Sigma_{15}$PASHs, concentrations of $\Sigma_4$PAOHs in non-harbour samples were *ca*. 10× smaller than the harbour samples.

There are a few studies that have reported on HPACs in Lake Ontario sediments. To identify coal tar markers, Zhang et al. analyzed carbazole, benzo[a]carbazole, dibenzothiophene and dibenzofuran in the bottom sediment samples from the same collection sites and similar concentration ranges were detected.[14] The presence of carbazoles and their halogenated derivatives in Lake Ontario sediments from the 1980s was reported earlier.[26] Also, dibenzothiophene and its alkylated derivatives have been reported as part of a fingerprinting analysis.[11]

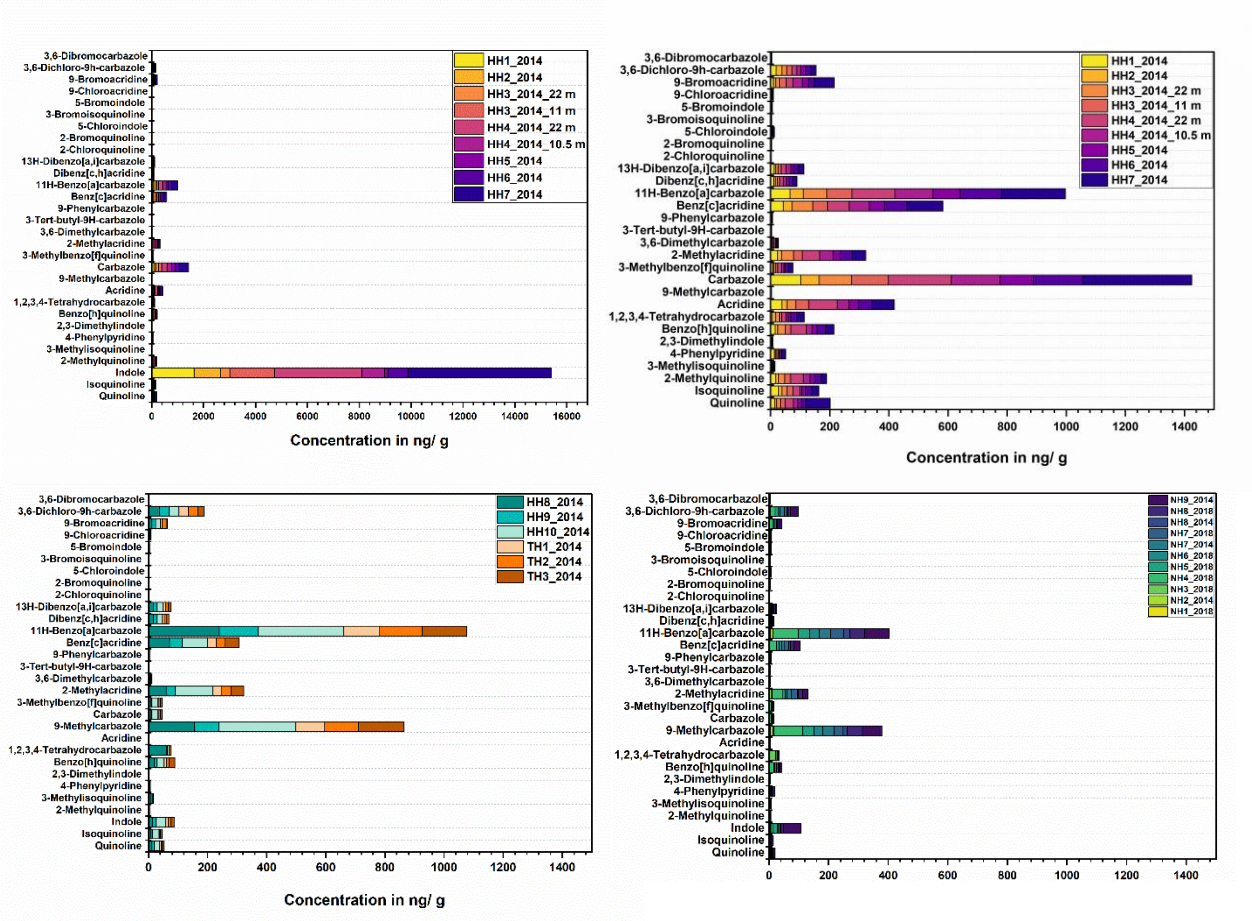



**Figure 4.1:** Concentrations of individual PANHs (ng/g, dry weight) in (a) the HH suspended sediments from 2014 with indole (top left panel), (b) without indole (top right panel), (c) the HH and TH bottom sediments from 2014 (bottom left panel) and (d) the non-harbour bottom sediments from 2014 and 2018 (bottom right panel).

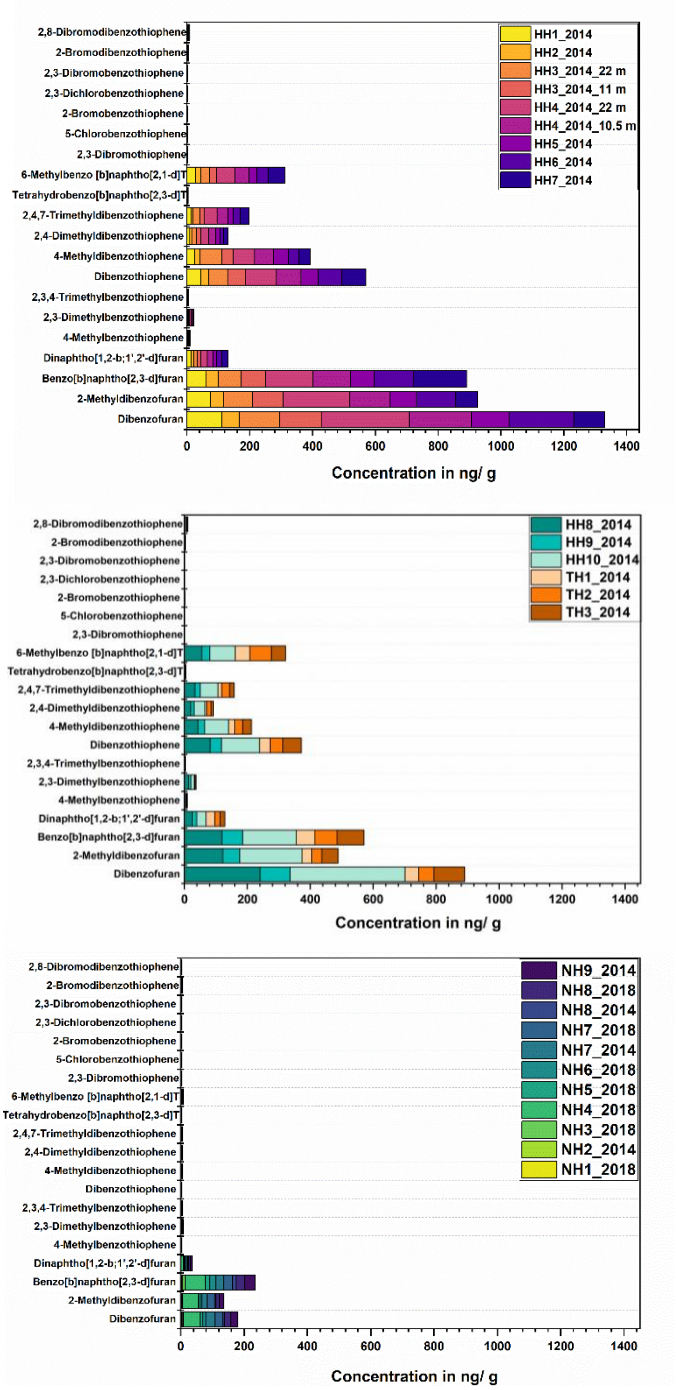



**Figure 4.2:** Concentrations of individual PASH and PAOH (ng/g, dry weight) in (a) the HH suspended sediments from 2014 (top panel), (b) the HH and TH bottom sediments from 2014 (middle panel) and (c) the non-harbour bottom sediments from 2014 and 2018 (bottom panel).

Figure II 2 summarises the temporal distribution of total PANH, PASH and PAOH concentrations in HH suspended sediments. Indole was purposely removed from the calculations of total PANHs. Based on the Mann-Kendall test, there were no significant temporal trends (p value > 0.05) for any analyte at any site over the 10-year study period.

### 4.4.2 Fingerprinting analysis of HPACs

A prerequisite for chemical fingerprinting in environmental media is that transformation processes between the source and sink are insignificant. Because the major loss mechanism in sediments is biodegradation, it is imperative to select chemicals known to be persistent in sediment for fingerprinting studies. Here we present two criteria we adopted to assign chemical persistence in suspended sediments and then the tools used for fingerprinting analysis.

The first criterion was based on a comparison of the chemical concentrations in suspended and bottom sediments. Hamilton Harbour site 9031 was used as a representative site for our assessment. If the concentration of the chemical in the bottom sediment exceeded that in the suspended sediment (see Figure II 3), then the chemical was considered persistent. The second criterion was based on our knowledge of the biodegradation of HPACs determined from empirical studies. Under controlled laboratory studies, Celik *et al*.[25] identified eight heterocyclic aromatic compounds that were persistent in adaptive microbial communities. Five of those persistent compounds and their structural isomers (benz[c]acridine, 11H-benzo[a]carbazole, dibenzothiophene, dibenzofuran and benzo[b]naphtho[2,3-d]furan) were measured in our study. Motivated by this, and with other biodegradation studies lacking, we used these five persistent compounds to assign persistence to other HPACs. Our approach was based on the premise that the concentrations of these five persistent HPACs will remain unchanged in the suspended sediments

over time. Temporal trend plots were first constructed using the log transformed concentration for each of the five persistent HPACs. Next, an average slope from these five plots was calculated along with an uncertainty based on a confidence interval (CI) at a 95% confidence level. Slopes for all the other HPACs in our analytical suite were derived in an identical manner, and persistence was ascribed if their respective slopes fell within the CI (Figure 4.3). The method of determining chemical persistence focuses solely on the magnitude of the slope and how it deviates from a defined value. However, it is important to recognize that a slope differing from 3.1 does not necessarily indicate a statistically significant trend in chemical concentrations over time. As a precautionary approach, chemicals needed to satisfy both criteria before they could be considered persistent. The chemicals that met both criteria and used for our chemical fingerprinting studies are shown in Table II 6. In total, twenty-five compounds met our persistence criteria. Of these, thirteen, eight and 4 were N-, S- and O-containing compounds, respectively. Most of the low-molecular weight two-ring PANHs and some of the three-ring systems (benzo[h]quinoline, acridine, 1,2,3,4-tetrahydrocarbazole and 9-bromoacridine) did not meet our persistence criteria. This agrees well with the known chemical properties of these compounds, i.e., higher water solubility, smaller octanol-water partition coefficients, smaller sorption constants, and, based on QSAR models, greater biodegradation potential (Table II 6). Admittedly, there were a few two-ring nitrogen and sulphur compounds that met our persistence criterion but, as a precaution, were not included in our fingerprinting analysis. Similarly, carbazole met our second persistence criteria (see Figure 4.3) however, its concentrations were smaller in bottom sediments compared to the suspended sediments and so was not considered persistent.[26] Biodegradation may contribute to the observed low concentrations of these heterocyclic aromatic compounds in the bottom sediment samples. These samples were collected from a depth range of 2–3 cm from the lake surface. In that

layer, oxygen and microbial activity are typically higher, which leads to a greater likelihood of aerobic microbial degradation of indole, its derivatives, and carbazole-like analytes.[27,28]

It should be acknowledged that there are some limitations with our approach. First, the lack of available field replicates precludes robust uncertainty measurements. Second, the impacts of sediment mixing and episodic input events into the lake were all considered negligible. Despite these limitations, the approach does provide a reasonable first approximation of chemical persistence in sediment in the absence of additional experimental biodegradation studies.

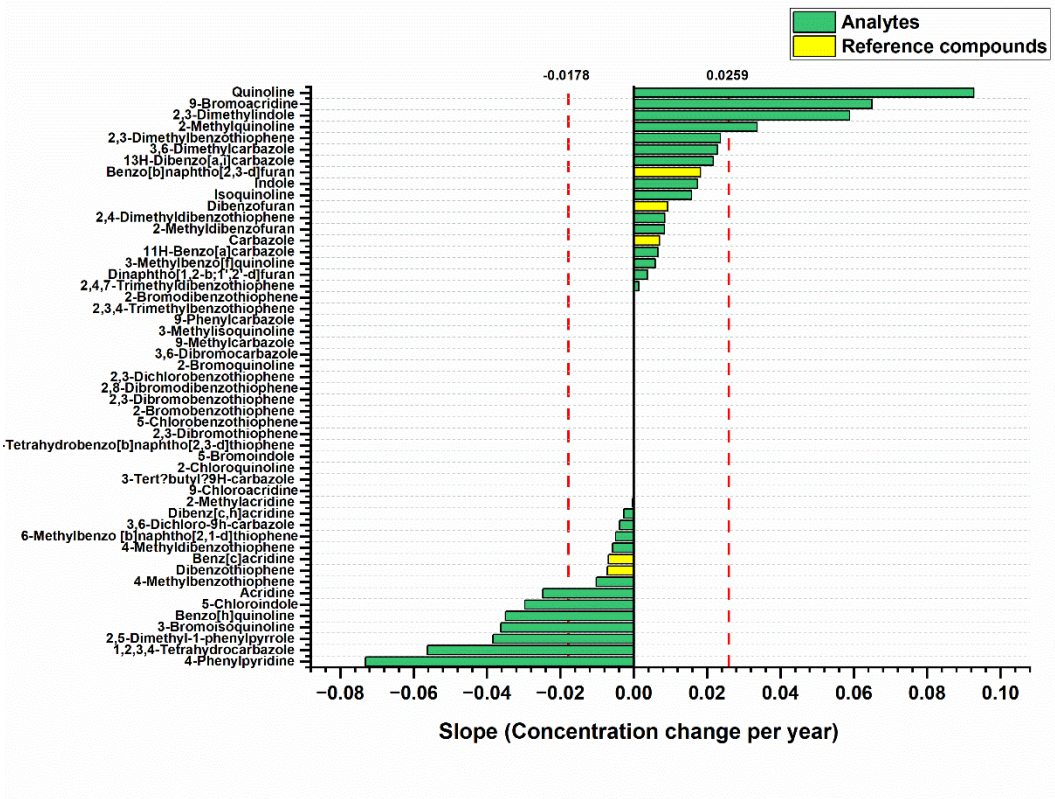



**Figure 4.3:** Evaluation of chemical persistence determined from the slopes (x-axis) of the log transformed concentration of each analyte vs sampling year for suspended sediments collected from Hamilton Harbour. The dotted red lines represent the confidence interval (at 95% confidence level) derived from the average slope of benz[c]acridine, 11H-benzo[a]carbazole, dibenzothiophene, dibenzofuran and benzo[b]naphtho[2,3-d]furan (shown in yellow).

#### 4.4.2.1 Diagnostic ratio

The ratios between the dibenzothiophene and its alkylated derivatives were calculated and compared between the HH suspended sediments to investigate potential sources of PACs. This analysis hypothesized that samples impacted by coal tar would exhibit lower degrees of alkylation than samples influenced primarily by lower combustion temperature processes, such as mobile emissions and urban runoffs. The DBT ratios calculated for the reference sample are depicted in Figure II 4, which lends support to our hypothesis that alkylation is lower in coal tar-based products compared to other non-coal tar-based reference materials such as tire rubber and asphalt-based products.

Dibenzothiophene ratios calculated for four different locations: station 914 located in tributary inflows, station 9083 located in industrial areas closer to Randle Reef remediation site, station 9032 located in open-water areas and station 9031 located near natural areas that have less direct impact by human activities. Figure II 5 summarizes DBT ratios calculated for these stations over the six time points. The PAC concentration data indicate that there is a discrimination between samples heavily impacted by coal tar-related sources *vs*. less impacted samples and those of mixed combustion source origin. As expected, the coal tar impacted samples from station 9083 showed the smallest ratios of the alkylated derivatives relative to the parent compound. Conversely, the ratios for station 914 (tributary inflow representing urban runoff from the watershed) show much larger ratios. The station at the west end of the harbour (9031) and the deep hole center station (9032) show intermediate profiles, likely because of a range of combustion process contributions, including legacy coal tar contamination and vehicular emissions.

#### 4.4.2.2 Compositional analysis

To identify the source of PACs using compositional analysis, the correlation between the total PAH profiles and typical combustion PAHs are evaluated. Idowu *et al*. used fluorene, pyrene, benz[a]anthracene, benzo[b]fluoranthene, benzo[k]fluoranthene and dibenz[a,h]anthracene as the combustion PAHs in their study.[29] In their study, to identify the relationship between the analytes and combustion PAHs, a Pearson correlation analysis was performed. The assumption of this analysis is that if there is a positive correlation between the analytes and the combustion PAHs (Pearson rho $\neq 0$, p-value $< 0$), the analytes are more likely to originate from combustion processes. Pearson correlation analysis requires the assumption that the data is normally distributed. To satisfy the assumption of normality, a Shapiro-Wilk normality test was performed on total concentrations of combustion PAHs, PANHs, PASHs, and PAOHs using RStudio. The p-values (p-value $< 0.05$) reject the null hypothesis that the data is normally distributed.

To further explore the correlation between the combustion PAHs and HPACs, a Spearman's correlation study was carried out. Unlike Pearson correlation analysis, Spearman's correlation analysis measures a monotonic relationship between variables without the assumption of linearity and normality. The Spearman's rank correlation rho values and approximated p-values for selected individual analytes and for the three different HPACs relative to total combustion PAHs are included in Table II 7. As expected, there were stronger statistical relationships between unsubstituted HMW and combustion PAHs than there were for the alkylated and halogenated derivatives and combustion PAHs.

There were strong positive statistically significant relationships for all PAC and total combustion PAHs (Figure SI6, $p<0.05$ in all cases). Based on the $R^2$-values, the rank order of the strength of

the relationships were ∑PANHs > ∑PASHs > ∑PAOHs. Taken together, the compositional analysis results suggest that these sites are contaminated with PACs that are more likely to originate from combustion-related processes.

#### 4.4.2.3 Multivariant analysis

Based on three lines of experimental evidence and prior to the multivariant analysis, reference materials were grouped into three categories: coal tar-based (CT RM), non-coal tar (NCT RM), and mixed sources (Mixed RM) (see Table II 8). As mentioned in the previous section, coal tar-based products have low alkylation, leading to low DBT ratios. Carbazole concentration and $\sum_{16}$PAH concentrations can also be used to delineate reference materials, as coal tar-based products have greater concentrations of these PACs compared to asphalt-based materials (see Table II 8).[30–32] Using this approach, four CT RM (polymer coal tar seal coat, coal tar emulsion seal coat, raw coal tar and railway ties pieces), seven NCT RM (tier rubber 1 & 2, asphalt-based sealer1, asphalt-based sealcoat2, asphalt shingle pieces, roof tar and crack filler liquid) and three Mixed RM (tar from tar & chip roadway (2015), airport grade driveway re-surfacer and non-fibrous roof coating) were identified.

Two dimensionality reduction, PCA and UMAP, were used to explore HPAC fingerprints in Lake Ontario sediment samples. PCA is widely used as a dimensionality reduction tool in chemical fingerprinting of environmental pollutants.[33–35] To ensure there was no autocorrelation between the temporal data, lag-1 Auto Correlation Factors (ACFs) were calculated using RStudio for all the analytes with respect to all the sites. No correlation was found between the temporal data, ensuring the validity of using all 69 samples in PCA. Out of the 25 persistent HPACs, 22 analytes had detection percentages above 10% and were used as the original variables of multivariate

analysis. PCA was performed for all the samples and reference materials, considering percent normalized concentrations of the selected 22 HPACs. PC1, PC2 and PC3 contributes 29.6%, 21.4% and 12.8% to the total variance, respectively. The low cumulative variance of the first three principal components (63.8%) indicates that the PCA fails to account for a large portion of the data. This can be explained by the large number of variables and the assumption of linearity between variables.

The loading values of each analyte for the first five principal components can be found in Table II 9 and the biplot is depicted in Figure II 7. The main analytes that contribute to PC1 are 7,8,9,10-tetrahydrobenzo[b]naphtho[2,3-d]thiophene and 2,3,4-trimethylbenzothiophene, benz[c]acridine, dibenzofuran and 2-methyldibenzofuran. High loading analytes for PC2 are 2,8-dibromodibenzothiophene, 9-chloroacridine, 3-tert-butyl-9H-carbazole, 2-methylacridine and 3,6-dichloro-9h-carbazole while dinaphtho[1,2-b;1',2'-d]furan, dibenzothiophene, 3-methylbenzo[f]quinoline, 9-methylcarbazole and dibenzofuran mainly contributes to PC3.

PCA plots represent the fingerprinting patters of for HH suspended sediments, HH bottom sediments, TH bottom sediments and NH bottom sediments and three groups of RMs the PCA score plots are depicted in Figure 4. The reference materials are dispersed into two main areas, along with PC1, indicating the distinct fingerprints between CT and NCT RMs. Despite some variations in PC2 and PC3, HH suspended sediment samples are clustered together and at the negative end of PC1 show some overlap with CT RMs, indicating the influence of legacy coal tar contamination in HH. HH bottom sediments shifted slightly towards to the center of the PCA space indicating reduced influence from CT. The bottom sediment samples from the two harbours show a mixed source characteristic. Non-harbour sediment samples showed a broader range in principle

components suggesting a diverse HPACs composition compared to harbour sediments and reference materials. Also, non-harbour sediments have high positive PC1 and PC2 scores compared to harbour sediments suggesting that these less industrialized or urbanized areas may have different HPACs sources such as atmospheric deposition and other natural sources. This was further corroborated as non-harbour sites such as Rouge River, Etobicoke Creek, Don River, Humber Bay near Humber River mouth and Humber River show a different fingerprinting pattern than harbour sediments and reference materials.

The assumption that a linear relationship between contaminants and their input to the sediments is likely inaccurate since these analyte concentrations can vary because of factors such as human activities, climate and geographical changes.[33] Therefore, a UMAP analysis was carried out as a non-linear dimensionality reduction technique. UMAP is known to give new dimensions with better structural preservation compared to PCA that can produce an artificial spread the data.[36]

UMAP analysis was carried out to produce three dimensions and the graph between the first two dimensions is shown in Figure 4.4 and this visualization shows more distinct groupings within the data that are not as apparent in PCA. For instance, there is a clear compositional variance between the two reference material types: CT and NCT. Like PCA, CT RM align more closely with suspended sediments supporting coal tar as a possible source. There is a very different chemical fingerprints between suspended and bottom sediments indicating vertical stratification and possible historical deposition in bottom sediments. The TH and NH samples serve as a clear background controls. The difference in HPACs fingerprints is likely due, in part, to the historical urban contamination of PACs in Hamilton Harbour. [37] When comparing the outputs of the two multivariant analysis, it is evident that PCA fails to capture local variability due to the strong

influence of background sediments. Meanwhile, UMAP plots reveal distinct clusters, demonstrating its ability to preserve local structures while generating outputs based on the selected hyperparameters.

Both PCA and UMAP highlighted that five reference materials (CT RM and one Mixed RM: airport grade driveway re-surfacer) were the main contributors to the chemical burden of HPACs in HH suspended sediments and showed that there were some differences between the HPAC fingerprints in suspended sediments and bottom sediments. Compared to bottom sediments, suspended sediments are more likely to represent recent and ongoing inputs of HPACs. Hence, to obtain a better resolution on the chemical fingerprints between HH suspended sediments and the reference materials, new PCA and UMAP analysis were conducted (Figure II 8) that excluded all three types of bottom sediments which reflects historical deposition and long-term accumulation patterns.

UMAP analysis on suspended sediments and reference materials further reveals two distinct chemical fingerprints between CT and NCT RM while the Mixed RM distributed between those two clusters indicate their shared origin. Even though there is only a small overlap between the suspended sediment samples and CT RM in the PCA plot, UMAP results clearly indicate CT RMs influenced HPACs fingerprints in the Hamilton Harbour. Closer analysis reveals some visible differences in chemical fingerprints in the samples from sites such as 9031 and Windermere 914 indicating more point sources compared to sites such as 9032 and 9033 where the data are more spread across the three variables. Also, UMAP was able to differentiate the similar fingerprints between the sediments from the known historical CT contaminated Randle Reef sites 9083 and 9081 and CT RMs with a very strong overlap of the two clusters compared to samples from other

sites. Taken together, our UMAP analysis confirms that all the suspended sediment samples have fingerprints comparable to coal tar.

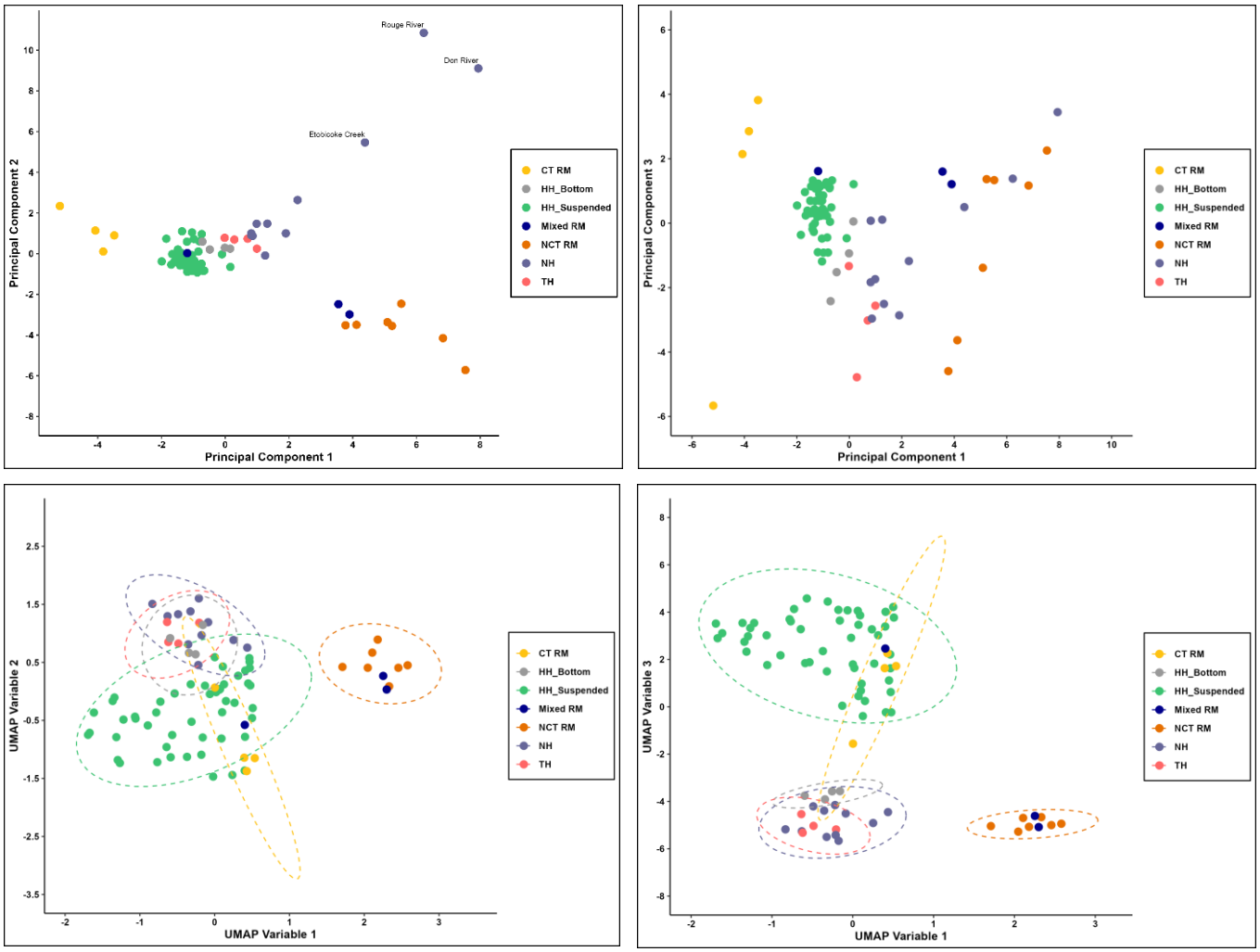



**Figure 4.4:** PAC and UMAP results between (a) PC1 and PC2 (top left panel), (b) PC1 and PC3 (top right panel) (c) variable 1 and variable 2 (bottom left panel), (d) variable 1 and variable 3 (bottom right panel) for HH harbour suspended sediments (green dots), HH harbour bottom sediments (gray dots), TH harbour Bottom sediments (red dots), non-harbour bottom sediments (purple dots), the CT reference materials (yellow dots), the NCT reference materials (orange dots) and the Mixed RM (dark blue dots).

## 4.5 CONCLUSIONS

A broad suite of heterocyclic aromatic compounds was detected in sediments from Lake Ontario which is a key geographical area in North America. The spatio-temporal trends suggest that the area is contaminated with N-, S and O-containing compounds, raising the question what impacts they might have on benthic organisms and animals higher up the food chain. New approaches to identify the environmental source of PACs to the region were investigated using HPACs data. To establish the prerequisites for the source attribution using fingerprinting analysis, an empirical approach based on environmental concentration differences and p-chem properties were used and out of fifty HPACs, twenty-five potential candidates were selected. Fourteen different reference materials were grouped into three categories: coal tar, non-coal tar and mixed based on the diagnostic ratios between the dibenzothiophene and its alkylated derivatives and concentration differences of carbazole and $\sum_{16}$PAHs. Compositional analysis suggests that the analytes of interests enter the environment predominantly from pyrogenic sources. Results of the multivariant analyses suggest that key contributors of HPACs in Hamilton Harbour are likely linked to coal tar from transport truck spills, railway tie fragments, polymer coal tar sealcoat, coal tar emulsion sealer, and airport-grade driveway resurfacing, while HPACs fingerprints of sediment samples collected from the Toronto Harbour and near-by non-harbour areas suggest that these areas are less impacted by coal tar based contaminants. The results confirm that heterocyclic aromatic compounds can be use as viable fingerprints for the source apportionment studies. Also, it further establishes the importance of incorporating novel fingerprinting techniques such as UMAP with more traditional approaches like PCA.

## 4.6 ACKNOWLEDGEMENTS

This work was supported by a Natural Sciences and Engineering Research Council Discovery Grant to GTT (Grant number: RGPIN-2024-04947).

## 4.7 REFERENCES


(1) Wallace, S. J.; de Solla, S. R.; Head, J. A.; Hodson, P. V.; Parrott, J. L.; Thomas, P. J.; Berthiaume, A.; Langlois, V. S. Polycyclic Aromatic Compounds (PACs) in the Canadian Environment: Exposure and Effects on Wildlife. *Environmental Pollution*. Elsevier Ltd October 1, **2020**. https://doi.org/10.1016/j.envpol.2020.114863.

(2) Sadiktsis, I.; de Oliveira Galvão, M. F.; Mustafa, M.; Toublanc, M.; Ünlü Endirlik, B.; Silvergren, S.; Johansson, C.; Dreij, K. A Yearlong Monitoring Campaign of Polycyclic Aromatic Compounds and Other Air Pollutants at Three Sites in Sweden: Source Identification, in Vitro Toxicity and Human Health Risk Assessment. *Chemosphere* **2023**, 332. https://doi.org/10.1016/j.chemosphere.2023.138862.

(3) Toxic substances list: schedule 1. Canadian Environmental Protection Act 1999, https://www.canada.ca/en/environment-climate-change/services/canadian-environmental-protection-act-registry/substances-list/toxic/schedule-1.html (Accessed on 3 January 2025).

(4) Andersson, J. T.; Achten, C. Time to Say Goodbye to the 16 EPA PAHs? Toward an Up-to-Date Use of PACs for Environmental Purposes. *Polycycl Aromat Compd* **2015**, 35 (2–4), 330–354. https://doi.org/10.1080/10406638.2014.991042.

(5) Marvin, C. H.; Tomy, G. T.; Thomas, P. J.; Holloway, A. C.; Sandau, C. D.; Idowu, I.; Xia, Z. Considerations for Prioritization of Polycyclic Aromatic Compounds as Environmental Contaminants. *Environmental Science and Technology*. American Chemical Society December 1, **2020**, pp 14787–14789. https://doi.org/10.1021/acs.est.0c04892.

(6) Ghosh, P.; Mukherji, S. Fate, Detection Technologies and Toxicity of Heterocyclic PAHs in the Aquatic and Soil Environments. *Science of the Total Environment*. Elsevier B.V. September 20, **2023**. https://doi.org/10.1016/j.scitotenv.2023.164499.

(7) Machado, M. E.; Nascimento, M. M.; Bomfim Bahia, P. V.; Martinez, S. T.; Bittencourt de Andrade, J. Analytical Advances and Challenges for the Determination of Heterocyclic Aromatic Compounds (NSO-HET) in Sediment: A Review. *TrAC - Trends in Analytical Chemistry.* Elsevier B.V. May 1, **2022**. https://doi.org/10.1016/j.trac.2022.116586.

(8) Abdel-Shafy, H. I.; Mansour, M. S. M. A Review on Polycyclic Aromatic Hydrocarbons: Source, Environmental Impact, Effect on Human Health and Remediation. *Egyptian Journal of Petroleum.* Egyptian Petroleum Research Institute March 1, **2016**, pp 107–123. https://doi.org/10.1016/j.ejpe.2015.03.011.

(9) Davis, E.; Walker, T. R.; Adams, M.; Willis, R.; Norris, G. A.; Henry, R. C. Source Apportionment of Polycyclic Aromatic Hydrocarbons (PAHs) in Small Craft Harbor (SCH) Surficial Sediments in Nova Scotia, Canada. *Science of the Total Environment* **2019**, 691, 528–537. https://doi.org/10.1016/j.scitotenv.2019.07.114.

(10) Idowu, I.; Francisco, O.; Thomas, P. J.; Johnson, W.; Marvin, C.; Stetefeld, J.; Tomy, G. T. Validation of a Simultaneous Method for Determining Polycyclic Aromatic Compounds and Alkylated Isomers in Biota. *Rapid Communications in Mass Spectrometry* **2018**, 32 (3), 277–287. https://doi.org/10.1002/rcm.8035.

(11) Marvin, C. H.; Mccarry, B. E.; Villella, J.; Allan, L. M.; Bryant, D. W. Chemical and Biological Profiles of Sediments as Indicators of Sources of Genotoxic Contamination in Hamilton Harbour. Part I: Analysis of Polycyclic Aromatic Hydrocarbons and Thia-Arene Compounds. *Chemosphere*, **2000**, 41(7), 979-988. https://doi.org/10.1016/S0045-6535(99)00494-4

(12) Sofowote, U. M.; Mccarry, B. E.; Marvin, C. H. Source Apportionment of PAH in Hamilton Harbour Suspended Sediments: Comparison of Two Factor Analysis Methods. *Environ Sci Technol* **2008**, 42 (16), 6007–6014. https://doi.org/10.1021/es800219z.

(13) Bowman, D. T.; Jobst, K. J.; Helm, P. A.; Kleywegt, S.; Diamond, M. L. Characterization of Polycyclic Aromatic Compounds in Commercial Pavement Sealcoat Products for Enhanced Source Apportionment. *Environ Sci Technol* **2019**, 53 (6), 3157–3165. https://doi.org/10.1021/acs.est.8b06779.

(14) Zhang, X.; Bowman, D. T.; Diamond, M. L.; Helm, P.; Jobst, K. J.; Hao, C.; Kleywegt, S.; Zhang, Z. F.; Marvin, C.; Zhang, X. Contribution of Coal Tar Sources to Polycyclic Aromatic Compounds and Associated Ecological Risk in Lake Ontario Sediments: Inference from a Novel Marker. *Environ Sci Technol* **2025**. https://doi.org/10.1021/acs.est.4c10360.

(15) Mari, M.; Harrison, R. M.; Schuhmacher, M.; Domingo, J. L.; Pongpiachan, S. Inferences over the Sources and Processes Affecting Polycyclic Aromatic Hydrocarbons in the Atmosphere Derived from Measured Data. *Science of the Total Environment* **2010**, 408 (11), 2387–2393. https://doi.org/10.1016/j.scitotenv.2010.01.054.

(16) Zou, Y.; Wang, L.; Christensen, E. R. Problems in the Fingerprints Based Polycyclic Aromatic Hydrocarbons Source Apportionment Analysis and a Practical Solution. *Environmental Pollution* **2015**, 205, 394–402. https://doi.org/10.1016/j.envpol.2015.05.029.

(17) Guo, J.; Li, Z.; Ranasinghe, P.; Bonina, S.; Hosseini, S.; Corcoran, M. B.; Smalley, C.; Rockne, K. J.; Sturchio, N. C.; Giesy, J. P.; Li, A. Spatial and Temporal Trends of Polyhalogenated Carbazoles in Sediments of Upper Great Lakes: Insights into Their Origin. *Environ Sci Technol* **2017**, 51 (1), 89–97. https://doi.org/10.1021/acs.est.6b06128.

(18) Wu, Y.; Tan, H.; Zhou, C.; Crimmins, B. S.; Holsen, T. M.; Chen, D. Bioaccumulation and Spatiotemporal Trends of Polyhalogenated Carbazoles in Great Lakes Fish from 2004 to 2016. *Environ Sci Technol* **2018**, 52 (8), 4536–4545. https://doi.org/10.1021/acs.est.8b00427.

(19) Burniston, D. A.; Jia, J.; Charlton, M. N.; Thiessen, L.; McCarry, B. E.; Marvin, C. H. Trends in Hamilton Harbour Suspended Sediment Quality. *Aquat Ecosyst Health Manag* **2016**, 19 (2), 141–149. https://doi.org/10.1080/14634988.2016.1172906.

(20) Vitharana, N. N.; Halldorson, T.; Xia, Z.; Francisco, O.; Marvin, C.; Thomas, P. J.; Liebzeit, E.; Lucas, A. M.; Moradi, V.; Tomy, G. T. A Validated Approach for Analysis of Heterocyclic Aromatic Compounds in Sediment Samples. *J Chromatogr A* **2024**, 1718. https://doi.org/10.1016/j.chroma.2024.464723.

(21) Xia, Z.; Idowu, I.; Kerr, E.; Klaassen, N.; Assi, H.; Bray, H.; Marvin, C.; Thomas, P. J.; Stetefeld, J.; Tomy, G. T. New Approaches to Reduce Sample Processing Times for the Determination of Polycyclic Aromatic Compounds in Environmental Samples. *Chemosphere* **2021**, 274. https://doi.org/10.1016/j.chemosphere.2021.129738.

(22) 'Dereviankin, M. Interactive GUI for tSNE and UMAP. GitHub (https://github.com/MikeDereviankin/Interactive_tSNE_UMAP).

(23) Mayer, T.; Bennie, D.; Rosa, F.; Palabrica, V.; Rekas, G.; Schachtschneider, J.; Marvin, C. Dispersal of Contaminants from Municipal Discharges as Evidenced from Sedimentary Records in a Great Lakes Coastal Wetland, Cootes Paradise, Ontario. *J Great Lakes Res* **2008**, 34 (3), 544–558. https://doi.org/10.3394/0380-1330(2008)34[544:DOCFMD]2.0.CO;2.

(24) Sullivan, D. W.; Gad, S. E. Indole. In Encyclopedia of Toxicology: Third Edition; Elsevier, **2014**; pp 1030–1031. https://doi.org/10.1016/B978-0-12-386454-3.00860-5.

(25) Çelik, G.; Stolte, S.; Müller, S.; Schattenberg, F.; Markiewicz, M. Environmental Persistence Assessment of Heterocyclic Polyaromatic Hydrocarbons – Ultimate and Primary Biodegradability Using Adapted and Non-Adapted Microbial Communities. *J Hazard Mater* **2023**, 460. https://doi.org/10.1016/j.jhazmat.2023.132370.

(26) Li, A.; Zhou, S.; He, H.; Guo, J.; Rockne, K. J.; Sturchio, N. C.; Giesy, J. P. Polyhalogenated Carbazoles in Sediments of Lower Laurentian Great Lakes and Regional Perspectives. ACS *ES and T Water* **2022**, 2 (9), 1544–1554. https://doi.org/10.1021/acsestwater.2c00191.

(27) Ma, Q.; Zhang, X.; Qu, Y. Biodegradation and Biotransformation of Indole: Advances and Perspectives. *Frontiers in Microbiology.* Frontiers Media S.A. November 1, **2018**. https://doi.org/10.3389/fmicb.2018.02625.

(28) Seo, J. S.; Keum, Y. S.; Li, Q. X. Bacterial Degradation of Aromatic Compounds. *International Journal of Environmental Research and Public Health.* MDPI **2009**, pp 278–309. https://doi.org/10.3390/ijerph6010278.

(29) Idowu, O.; Carbery, M.; O'Connor, W.; Thavamani, P. Speciation and Source Apportionment of Polycyclic Aromatic Compounds (PACs) in Sediments of the Largest Salt Water Lake of Australia. *Chemosphere* **2020**, 246. https://doi.org/10.1016/j.chemosphere.2019.125779.

(30) Mahler, B. J.; Van Metre, P. C.; Foreman, W. T. Concentrations of Polycyclic Aromatic Hydrocarbons (PAHs) and Azaarenes in Runoff from Coal-Tar- and Asphalt-Sealcoated Pavement. *Environmental Pollution* **2014**, 188, 81–87. https://doi.org/10.1016/j.envpol.2014.01.008.

(31) Ahrens, M. J.; Depree, C. V. A Source Mixing Model to Apportion PAHs from Coal Tar and Asphalt Binders in Street Pavements and Urban Aquatic Sediments. *Chemosphere* **2010**, 81 (11), 1526–1535. https://doi.org/10.1016/j.chemosphere.2010.08.030.

(32) Germin-Aizac, J.; Maitre, A.; Balducci, F.; Montlevier, S.; Marques, M.; Tribouiller, J.; Demeilliers, C.; Persoons, R. Bitumen Fumes and PAHs in Asphalt Road Paving: Emission Characteristics, Determinants of Exposure and Environmental Impact. *Environ Res* **2023**, 228. https://doi.org/10.1016/j.envres.2023.115824.

(33) Zhou, M.; Li, Y. Spatial Distribution and Source Identification of Potentially Toxic Elements in Yellow River Delta Soils, China: An Interpretable Machine-Learning Approach. *Science of the Total Environment* **2024**, 912. https://doi.org/10.1016/j.scitotenv.2023.169092.

(34) Dvorská, A.; Komprdová, K.; Lammel, G.; Klánová, J.; Plachá, H. Polycyclic Aromatic Hydrocarbons in Background Air in Central Europe - Seasonal Levels and Limitations for Source Apportionment. *Atmos Environ* **2012**, 46, 147–154. https://doi.org/10.1016/j.atmosenv.2011.10.007.

(35) Liu, Y.; Chen, L.; Huang, Q. hui; Li, W. ying; Tang, Y. jian; Zhao, J. fu. Source Apportionment of Polycyclic Aromatic Hydrocarbons (PAHs) in Surface Sediments of the Huangpu River, Shanghai, China. *Science of the Total Environment* **2009**, 407 (8), 2931–2938. https://doi.org/10.1016/j.scitotenv.2008.12.046.

(36) Russo, A.; Borras, A. Comparison of Dimension Reduction Techniques Applied to the Analysis of Airborne Radionuclide Activity Concentration. *J Environ Radioact* **2022**, 244–245. https://doi.org/10.1016/j.jenvrad.2022.106813.

(37) Marvin, C. H.; Berthiaume, A.; Burniston, D. A.; Chibwe, L.; Dove, A.; Evans, M.; Hewitt, L. M.; Hodson, P. V.; Muir, D. C. G.; Parrott, J.; Thomas, P. J.; Tomy, G. T. Polycyclic Aromatic Compounds in the Canadian Environment: Aquatic and Terrestrial Environments. *Environmental Pollution*. Elsevier Ltd September 15, **2021**. https://doi.org/10.1016/j.envpol.2021.117442.

# CHAPTER 5 USING CHEMOMETRICS TO DISCERN PATTERNS AND SPATIAL DISTRIBUTION OF HETEROCYCLIC AROMATIC COMPOUNDS IN LAKE ERIE FISH

In this chapter, I present my work on the detection of HPACs in biota from three basins of Lake Erie (eastern, central, and western). Using advanced chemometric tools, I was able to show that there are differences in PAC profiles among the fish species analysed. I was also able to show that HPACs show a more distinct spatial dependency than PAHs and APAHs, further illustrating the advantage of HPACs for source identification.

A version of this chapter has already been published as Vitharana Nipuni N., Halldorson Thor, Paddock Brett, Tomy Thane, Gerring Aviva, Marvin Chris, Heuvel C. E., Fisk A.T., Tomy Gregg T., Using Chemometrics to Discern Patterns and Spatial Distribution of Heterocyclic Aromatic Compounds in Lake Erie Fish, Environmental Pollution, 2026 Copyright © 2026. B.V. Elsevier. Reprinted with permission.

## 5.1 ABSTRACT

Polycyclic aromatic compounds (PACs) are ubiquitous environmental contaminants, and while polycyclic aromatic hydrocarbons (PAHs) are well-studied, their heterocyclic counterparts (HPACs) remain poorly understood. We examined the spatial distribution and bioaccumulation of 104 PACs in four fish species [freshwater drum (*Aplodinotus grunniens*), yellow perch (*Perca flavescens*), gizzard shad (*Dorosoma cepedianum*), and walleye (*Sander vitreus*)] from the three main basins of Lake Erie, North America. Total (∑) PAC concentrations ranged from 0.7 to 43.2 μg/g lipid weight, with the highest levels found in biota from the western basin. Regardless of the basin and species type, HPACs were the most abundant contaminants, representing over 50% of $\sum_{104}$PACs in most samples. New and existing chemometric approaches revealed distinct-basin and species-dependent clustering in PAC profiles. Hierarchical cluster analysis highlighted enrichment of brominated HPACs in the central basin, particularly its western region, with smaller levels in top predator walleye, suggesting limited trophic transfer. Overall, the identification of previously unreported HPACs in Great Lakes biota, coupled with advanced chemometric approaches to resolve spatial and species-specific patterns, improves our understanding of PAC bioaccumulation and distribution in freshwater ecosystems and supports ecological risk assessments aimed at protecting the Great Lakes.

## 5.2. INTRODUCTION

Polycyclic aromatic compounds (PACs) are a group of organic contaminants that originate from various sources, including the combustion of organic matter, petroleum-based activities, and natural processes.[1] Due to their complex structures, these compounds are classified into different subclasses such as polycyclic aromatic hydrocarbons (PAHs), alkylated polycyclic aromatic hydrocarbons (APAHs), halogenated polycyclic aromatic hydrocarbons (HPAHs) and hetero-polycyclic aromatic compounds (HPACs).[2] For many years, scientific research has primarily concentrated on PAHs because of their known toxic effects and widespread occurrence.[3–5] However, more recent studies have shown that HPAHs and HPACs can exhibit similar or even greater toxicity than PAHs, highlighting the need to better understand their behavior and impact on the aquatic environment.[6,7]

Hetero-polycyclic aromatic compounds contain at least one nitrogen, sulfur, and/ or oxygen-like heteroatom in the aromatic system, which leads to having higher polarity compared to their homocyclic analogous compounds. This increased polarity enhances their water solubility and bioavailability in aquatic environments.[8] Despite their higher bioavailability, research on the fate and behavior of HPACs in biota has been limited due to several analytical challenges. These challenges include the high structural diversity and complexity of these compounds and a lack of commercially available standards and certified reference materials.[9]

PACs exist as complex mixtures, with environmental inputs arising from multiple sources, including both natural and anthropogenic origins.[10] A prerequisite for reliable source attribution is chemical persistence in the studied environmental compartment.[11] While fingerprinting methods are well-established, their application to biological compartments is complicated by

biotransformation, which reduces compound persistence compared to abiotic systems. Given the current paucity of data regarding PAC persistence in biota, meaningful source apportionment studies are not feasible.

The Laurentian Great Lakes, as the world's largest freshwater system, are of great importance for their rich biodiversity, economic activities like fisheries, and recreation.[12] Despite being the shallowest and smallest of the Great Lakes by volume, Lake Erie's fisheries are especially vital to the local economy.[13] The lake is also unique, as it is naturally divided into three distinct basins - Western, Central, and Eastern – based on differences in depth and water characteristics. These ecological variations across the basins also lead to variation in fish communities in each area.[14]

Historically, Lake Erie has been subject to contamination from various organic pollutants. Prior studies have focused on the bioaccumulation of different legacy contaminants in fish populations.[15] While some historical research has investigated PAHs in this region, there is a notable absence of data on both APAHs and HPACs.[16] Furthermore, recent data on PAH concentrations in the lake is limited. While some research has been conducted on HPAC concentrations in sediments from the Great Lakes, the available data for biota is significantly less.[17–19] To our knowledge, no previous studies have specifically focused on understanding the distribution of HPACs in biota from any of the Great Lakes.

Based on our recent findings on the distribution of HPACs and their halogenated derivatives in sediment and their potential for high bioavailability, we hypothesized that Lake Erie fish are contaminated with these compounds. Given their different chemical and physical characteristics, we expected the concentrations and patterns to vary in fish based on their feeding ecology and across the lake, given variation in the limnological characteristics of the basins.[20] To test this

hypothesis, we developed and validated a new extraction and purification method to analyze a broad suite of PACs in biological samples. Four fish species [freshwater drum (*Aplodinotus grunniens*), yellow perch (*Perca flavescens*), gizzard shad (*Dorosoma cepedianum*), and walleye (Sander vitreus)] with different feeding ecologies were collected from the eastern, central, and western basins of Lake Erie. We also used existing and newer chemometric methods to further explore our information-rich dataset to help discern distribution patterns of PACs in biota from different geographical regions of the lake.

## 5.3. MATERIALS AND METHODS

### 5.3.1 Chemicals

All PAHs, APAHs and labelled PAHs were purchased from Chiron (AS Stiklestadveien 1, 7041 Trondheim, Norway). The HPACs were purchased from Chiron and Sigma Aldrich Canada Ltd (Oakville, Ontario, Canada). High-purity (Optima grade) organic solvents, silica gel (923 grade, 100 - 200 mesh), alumina (60 - 325 mesh), anhydrous sodium sulfate and diatomaceous earth (DE) were purchased from Fisher Chemicals (Ottawa, Canada). The targeted analytes and internal standards, their codes and their CAS numbers can be found in Table III 1 and III 2. For recovery internal standard (RIS) naphthalene-$d_8$, fluorene-$d_{10}$, phenanthrene-$d_{10}$, pyrene-$d_{10}$, benz[a]anthracene-$d_{12}$, chrysene-$d_{12}$, benzo[b]fluoranthene-$d_{12}$, benzo[a]pyrene-$d_{12}$ and benzo[g,h,i]perylene-$d_{14}$ were used. Labeled anthracene (anthracene-$d_{10}$) was used as the instrument performance internal standard (IPIS).

### 5.3.2 Biota samples

Four different fish species were collected from Lake Erie from May to October 2019. The sample sites are depicted in Figure III 1. The protocols followed during sample collection and storage are described in Heuvel et al.[20] and Appendix III.

### 5.3.3 Extraction and cleanup procedures

A new extraction and purification procedure was developed and validated for simultaneous analysis of polycyclic aromatic nitrogen heterocycles (PANHs), polycyclic aromatic sulfur heterocycles (PASHs), and polycyclic aromatic oxygen heterocycles (PAOHs). Approximately 5 g of wet fish muscle tissue samples were dried using a LABCONCO freeze drier for 48h, and then mixed with 3.5 g of baked DE dispersant and added to an accelerated solvent extraction (ASE) cell fitted with a fiber glass filter at the bottom of the cell. Each cell was spiked then with 10.0 µL of 1 ng/µL mixture of RIS and filled with baked Ottawa sand to fill the void volume. Average RIS recoveries are summarised in Table III 3. Samples were extracted using a 50:50 mixture of DCM: ethyl acetate using a Dionex Accelerated Solvent Extractor (ASE) 350 (ThermoFisher, Mississauga, ON). The extraction was carried out with system pressure of 1500 psi, oven temperature of 120 $^{\circ}$C, static time of 15 min with 2 static cycles, flush volume of 60% and a nitrogen purge for 80 s at 150 psi. Samples were dried by directly adding sodium sulfate to the sample extracts and filtering through glass wool. Then volume was reduced to 5 mL and 0.5 mL was aliquoted for gravimetric lipid weight determination (Table III 4). The remaining sample volume was applied to the head of a gel permeation chromatography (GPC) column (J2-scientific AccuPrepMPSTM, Columbia, Missouri, USA) packed with 60 g of S-X3 Biobeads using 50:50

(v/v) DCM: hexane as the mobile phase with the pressure of 8 psi and flow rate of 5 mL/min. The GPC fractions were determined by spiking native analytical standards into 1 g of cod liver oil and running through the GPC column using the above solvent system. The first 120 mL containing lipids and other matrices was discarded, and the fraction between 120 mL and 250 mL was collected for further processing. Then the sample volume was reduced to 2 mL using a rotor evaporator. Then a mixture of 4 g of silica gel and 1 g of 5 % deactivated alumina and 0.5 g of sodium sulfate was prewashed with 10 mL of 70:30 (v/v) DCM: hexane and then 10 mL of ethyl acetate and transferred to the flat bottom flask containing the sample with 2 x 10 mL of 70:30 (v/v) DCM: hexane mixture. Samples were swirled intermittently for 30 minutes and then transferred to 60 mL test tubes via glass wool with 2 x 5 mL of 70:30 (v/v) DCM: hexane rinse and 10 mL of ethyl acetate rinse. Sample volume was reduced to 100 µL using a gentle stream of ultra-high purity (UHP) nitrogen, spiked with 10.0 µL of 1 ng/µL of IPIS (anthracene-$d_{10}$) prior to the GC-MS/MS injection.

### 5.3.4 Gas chromatography tandem mass spectrometry

An Agilent 7890 GC coupled with a triple quadrupole Agilent 7000C mass spectrometer (GC-MS/MS) with an electron ionization (EI) source was used for the targeted analysis of all the analytes. An Agilent J&W DB-5ms ultra inert column (30 m x 0.25 mm, 0.25 µm) and helium was used as the carrier gas at a constant flow rate of 1.2 mL/min. Split/splitless injections of 1 µL were achieved using a PAL RSI 85 auto sampler into our injector port, maintained at 250°C. Further information on the multiple reaction monitoring (MRM) methods used is described by Vitharana *et al*.[21] and Idowu *et al*.[22]

### 5.3.5 Method validation

The developed extraction and purification method was validated according to the Eurachem Guide -The Fitness for the Purpose of Analytical Methods.[23] Working range of HPACs for the GC/MS-MS was previously determind.[21] Accuracy, repeatability, intermediate precision, detection limits and ruggedness was checked with matrix blanks fortified with three concentration levels (25, 125 and 500 pg/µL). Due to the unavailability of commercial standard reference materials store-bought wild salmon sample was homogenised, freeze dried and used as the matrix. Six matrix blanks were processed without spiking the HPACs as a quality control measure. Complete steps followed during the method validation and the performance characteristics results for HPACs (Table III 5) can be found in Appendix III. Method validation parameters for homocyclic PACs were determined previously.[22]

To maintain our ISO 17027:2018 requirements, our laboratory participates twice annually in the Wepal-Quasimeme (Netherlands) proficiency testing program. For all 33 PACs (PAHs, APAHs and dibenzothiophenes) quantified in a biota certified reference material, z-scores were consistently < 2, confirming that our method is fit for its intended purpose for these analytes. To our knowledge, there are no PT exercises that capture the entire suite of analytes measured in this study.

### 5.3.6 Statistical analysis

The statistical analysis was performed using RStudio.[24,25] To explore the variations between the PAC profiles in the different basins, concentrations of each PAC class across those were compared with a series of statistical tests. The lipid corrected concentration data were log transformed prior to the statistical tests to improve the normality and homogeneity of variances. Then, normality and

homogeneity were assessed using the Shapiro–Wilk and Levene's tests, respectively. When both assumptions were satisfied, differences among basins were evaluated using one-way ANOVA, followed by Tukey's honestly significant difference (HSD) test for pairwise comparisons. When normality was met, but variances were unequal, Welch's ANOVA was applied, followed by Games–Howell post hoc tests to identify the PAC. For datasets that did not meet normality assumptions after transformation, non-parametric Kruskal–Wallis tests were performed on the log transformed concentration data, followed by Dunn's post hoc test. Statistical significance was defined as $p < 0.05$.

To further identify the HPAC variations between the four basins, PCA, UMAP, and HCA were performed. The concentration values less than the lower detection limit (LOD) were replaced by 0.01 ng/g to avoid the distortions caused by zero. By creating a binary table, the detects and non-detects were identified. For each sample, the percent detection of each analyte (detection percentage) was determined and the analytes ≥ 75% were selected for PCA and UMAP analysis. Then, to minimize the species-dependent variations, the data were log transformed, and one-way ANOVA was performed considering analyte-basin combinations. The analytes with nonsignificant variations (p value < 0.05) in at least three basins were selected for PCA and UMAP analysis. Prior to the multivariant analysis the original lipid corrected concentration data of the selected analytes was Hellinger transformed by taking the square root of the proportions between the concentrations and the total PAC concentrations of that specific sample. This reduces the dominance of highly abundant variables and increases the influence of low concentrated analytes in cluster formation, allowing Euclidean-based methods to capture subtle patterns in low-abundance analytes without the skew associated with raw concentration data. To explore the local structure of the data, the

hyperparameters used for UMAP analysis were n_neighbors = 5, min_dist = 0.01, and n_components = 3.

Analytes that had detection percentages > 90% were selected for the HCA. The Hellinger transformed data was used to compute Bray-Curtis dissimilarities, and hierarchical clustering was performed using the average linkage method. The HCA was performed considering all species and individual species to create heatmaps to identify the HPAC profiles between Lake Erie basins.

## 5.4. RESULTS AND DISCUSSION

### 5.4.1. Profiles and concentrations of polycyclic aromatic compounds in biota

Due to statistical differences in mean concentrations of several compounds (p-value of Student t-test < 0.05, Table III 6), comparatively high % RSTD values, and the greater spread in the geographical locations of collection sites, the central basin samples were further divided into two groups, i.e., west central and east central (Figure 5.1). This was done to ensure a better representation of the spatial distribution of PACs in fish.

Table 5.1 summarizes the arithmetic mean (± SD), median, and range of the lipid normalized concentrations of ∑31PANHs, ∑20PASHs, ∑4PAOHs, ∑16PAHs, and $\sum_{33}$APAHs detected in the four fish species collected from the four regions of Lake Erie. To account for non-normal data, the log transformed concentrations (ng/g, lw) are visually depicted in Figure 5.1. In general, and irrespective of sampling location, the rank order of the median PAC concentrations in biota was $\sum_{31}$PANHs > $\sum_{37}$APAHs > $\sum_{16}$PAHs > $\sum_{20}$PASHs > $\sum_{4}$PAOHs.[8]

Figure III 2 to III 5 shows the concentration profiles of each of the different PAC types; however, for simplicity and because concentrations were highest in fish from the western basin, we will

restrict our discussion to chemical profiles in fish from this location. Indole and its halogenated derivatives and 4-phenylpyridine contribute prominently to the $\sum_{31}$PANH burdens in fish from this basin. For freshwater drum, indole (median: 0.5 µg/g, lw), 5-chloroindole (median: 1.1 µg/g, lw) and 5-bromoindole (median: 1.1 µg/g, lw) account for 15, 30 and 30%, respectively, of the $\sum_{31}$PANH. In all other fish, concentrations of indole were greater than its halogenated derivatives. In yellow perch, for example, indole (median: 7.4 µg/g, lw) accounted for 87% of the $\sum_{31}$PANH, while halogenated indoles accounted for < 14%.

Inputs of indole into the aquatic environment and its derivatives can be both natural and anthropogenic. Bacterial conversion of tryptophan to indole by tryptophanase is one of the major routes of indole production.[26]. Anthropogenic inputs of indole primarily arise from coal tar-based industries, human and animal waste, and tobacco smoke.[27] These combined sources likely account for the relatively high abundance of indoles in biological systems compared with other PANHs. Indole and its derivatives are known AhR activators that have also shown toxic properties such as hemoglobinuric nephrosis, skin irritations, and tumor formation.[28,29]

Other two-ring PANHs, such as quinoline and isoquinoline, and three-ring PANHs, such as carbazole and acridine, were detected in all fish species at higher concentrations than their alkylated derivatives and high molecular weight (HMW) PANHs. The relationship between physicochemical properties (e.g., Kow) and HPAC accumulation remains difficult to fully evaluate, as there is a notable absence of controlled lab-based studies investigating the biotransformation rates of these heterocyclic compounds. While higher Kow typically correlates with higher accumulation in recalcitrant compounds, the lower-than-predicted concentrations of HMW compounds observed in our field samples likely reflect efficient metabolic processing, a

phenomenon well-documented for other polycyclic structures. Conversely, the prevalence of low molecular weight (LMW) compounds suggests that their accumulation may be driven by continuous environmental exposure rather than dietary magnification alone.

Interestingly, the arithmetic mean and median concentrations of 2-bromoisoquinoline, and 9-bromoacridine were greater in freshwater drum compared to the other three species, whereas the unsubstituted PANHs were more abundant in those species relative to their halogenated derivatives. Freshwater drum are benthic feeders, suggesting halogenated PANHs in higher concentrations in this species could be more associated with sediments.

Another prominent PANH detected in biota from the western basin, especially in yellow perch and gizzard shad, was 4-phenylpyridine. The lipid corrected median concentration of 4-phenylpyridine in yellow perch and gizzard shad were 1.2 μg/g and 1.2 μg/g, which accounted for 12 and 29% of the total PANH burden, respectively. These were higher than concentrations in walleye (median: 0.3 μg/g, lw) and in freshwater drum (median: 0.07 μg/g, lw). The differences in concentrations between the species are difficult to understand, especially with limited data on its distribution in water, sediment, and lower trophic levels. Yellow perch and gizzard shad are known zooplankton feeders,[30] which may be a source of 4-phenylpyridine. The lower concentrations in the walleye, a higher trophic level species[31] suggest trophic transfer is not high and the chemical does not biomagnify. Freshwater drum feed benthically[32] and the low concentrations suggest 4-phenylpyridine is not accumulated from sediments. This chemical is used as a building block in the synthesis of pharmaceuticals and organic light-emitting diodes.[33,34] Improper disposal of industrial waste or run-off from manufacturer sites are possible point sources of 4-phenylpyridine

into the environment, and there is a need to better understand it's distribution throughout the food web and in abiotic components.

Total mean and median concentrations of carbazoles in yellow perch and walleye from the western basin were both dominated by 1,2,3,4-tetrahydrocarbazole. In yellow perch, the total median carbazole concentration (sum of 10 compounds) was 0.1 μg/g (lw), with 1,2,3,4-tetrahydrocarbazole (median: 0.06 μg/g, lw) accounting for *ca*. 67% of the total. In walleye, the median concentration of 1,2,3,4-tetrahydrocarbazole was 0.4 μg/g (lw), which accounts for *ca*. 93% of the total carbazole burden. While unsubstituted carbazole was the most abundant carbazole derivative in gizzard shad (median: 0.01 μg/g, lw) from the western basin, 1,2,3,4-tetrahydrocarbazole contributed *ca*. 30% of the total carbazole burden. The most abundant carbazole in western basin freshwater drum samples, 3,6-dichlorocarbazole, had a median concentration of 0.1 μg/g (lw), accounting for 61% of total carbazoles, while all other carbazoles were an order of magnitude smaller.

In 2018, Wu *et al*.[35] investigated the distribution of polyhalogenated carbazoles in fish collected from all five Great Lakes. The total concentration of the 11 halogenated carbazoles reported was 66.1 ± 20.3 ng/g (lw) in lake trout (*Salvelinus namaycush*) and walleye collected from Lake Erie. In comparison, concentrations of total carbazoles in Great Lakes fish ranged from 54.7 to 154 ng/g (lw). Interestingly, fish from Lake Erie had the lowest concentrations of chlorinated carbazoles compared to all the Great Lakes. This indicates a higher prevalence of these compounds in other lakes and underscores the need for continued and widespread monitoring of PACs in Great Lakes areas.

Compared to PANHs, PASHs and PAOHs were detected at much smaller concentrations across all species. The highest mean/ median concentration of $\sum_{20}$PASHs was detected in freshwater drum samples collected from the western basin [median: 4.0 μg/g (lw)]. The 20 PASHs reported here include alkylated C1-C4 dibenzothiophenes that were quantified using the cluster method commonly used in APAH analysis.[22] The highest concentration of alkylated dibenzothiophenes and halogenated PASHs was in freshwater drum from the western basin (3.7 μg/g and μg/g, lw, respectively), accounting for 92 and 1.3%, respectively, of the total PASHs. Lipid corrected concentrations increased with the degree of alkylation, consistent with greater relative bioaccumulation potential of higher alkylated congeners.

Based on these results, there were several differences between HPAC concentrations in freshwater drum and the other fish species. Freshwater drum has high concentrations of halogenated PANHs, halogenated PASHs and alkylated dibenzothiophenes. With the increased nonpolar characteristics introduced with halogenation and alkylation, these HPACs have a higher tendency to bind to sediments. As a result, freshwater drum, a benthic and low trophic level species[32] is exposed to a greater extent to these compounds relative to the pelagic species such as gizzard shad, yellow perch and walleye.

Out of the four PAOHs, dibenzofuran was the dominant species across all four species. The median concentration of dibenzofuran in both yellow perch and the gizzard shad was ca. 0.06 μg/g (lw), accounting for >75% of $\sum_{4}$PAOHs. In general, concentrations of PAOHs decreased with increasing molecular weight across all four species. Walleye had the lowest PAOH burden with a median $\sum_{4}$PAOHs concentration of 0.01 μg/g (lw). However, concentrations of PAOHs exhibited high variability among individual samples, with percent relative standard deviation (% RSD)

exceeding 100% in all cases. The high degree of variability suggests heterogeneous exposure sources or species-specific differences in uptake and metabolism of these compounds.

Compared to total heterocyclic aromatic compounds, total PAH and APAH concentrations were consistently lower or similar across all species. Records of PAHs in Lake Erie sediments and biota date back to the 1980s.[16] Well-established PAC-generating processes, such as crude oil–related activities, pyrogenic sources, and coal tar–based industries, typically release a higher proportion of homocyclic aromatic compounds into the environment relative to HPACs.[4,36,37] For instance, crude oil contains 0.1–2.0% total nitrogen, of which less than 30% can be attributed to aromatic PANHs.[38] Consequently, it is generally expected that PAHs and APAHs would be more prevalent in environmental systems than HPACs, a trend also observed in sediment samples from Lake Ontario analyzed for the same suite of compounds.[37] However, PAC profiles from the present study indicate that, in Lake Erie fish muscle, HPACs were more abundant than their homocyclic counterparts. The lower concentrations of PAHs and APAHs compared to HPACs likely reflect their reduced bioavailability, lower trophic transfer and bioaccumulation potential due to greater biotransformation in biota.

Across all samples two of the most frequently detected PAHs in fish were fluorene (median: 0.02 to 0.1 μg/g, lw) and phenanthrene (median: 0.06 to 0.2 μg/g, lw). Fluoranthene and pyrene were also present at elevated concentrations in certain samples, e.g., yellow perch and gizzard shad from the central basin. Among the 33 APAHs analyzed, C2-chrysene was consistently detected at elevated levels in all species, independent of geographic location. Alkylated naphthalenes and alkylated fluorenes were also prominent analytes across all four species. APAH concentrations

were notably higher in yellow perch and freshwater drum collected from the west-central and western basins compared to other species and locations.

Unlike other PAC types, there are several reports documenting the presence of PAHs in Lake Erie biota. In 1997, the $\sum_{16}$PAHs in zebra mussels from the western basin ranged from 790 ± 370 to 1840 ± 1466 ng/g (lw).[39] In 2000, Gewurtz *et al*. reported PAH concentrations in four different benthic species (mayfly larvae, dreissenia, amphipods, and crayfish) from the western basin with individual concentrations ranging from 0.02 ± 0.01 to 7.60 ± 0.43 μg/g (lw).[40] Despite differences in species between past studies and this one, the PAH concentrations in Lake Erie biota remain comparable over two decades, suggesting a relative environmental steady state for PAHs.

**Table 5.1:** Arithmetic mean (± SD), median and range of the lipid normalized concentrations (µg/g) of different PAC groups detected in the four fish species collected from the four regions of Lake Erie

| Species | Basin | $\sum_{31}$PANH | $\sum_{20}$PASH | $\sum_{4}$PAOH | $\sum_{16}$PAH | $\sum_{33}$APAH | $\sum_{105}$PAC |
|---|---|---|---|---|---|---|---|
| **Freshwater Drum** | Eastern (n = 5) | 6.5 ± 8.1 | 0.2 ± 0.1 | 0.2 ± 0.1 | 0.08 ± 0.07 | 0.9 ± 0.5 | 7.8 ± 8.6 |
| | | 4.4 | 0.03 | 0.3 | 0.07 | 0.8 | 5.4 |
| | | 0.2 – 20.0 | 0.04 - 0.3 | 0.02 – 0.4 | 0.01 - 0.2 | 0.2 – 1.6 | 0.8 - 22.3 |
| | East Central (n = 2) | 0.1 ± 0.7 | 0.1 ± 0.02 | 0.008 ± 0.003 | 0.1 ± 0.1 | 0.9 ± 0.2 | 1.4 ± 0.5 |
| | | 0.1 | 0.02 | 0.008 | 0.1 | 0.9 | 1.7 |
| | | 0.09 – 0.2 | 0.01 - 0.03 | 0.006 - 0.01 | 0.09 – 0.2 | 0.9 – 1.1 | 1.0 - 1.7 |
| | West Central (n = 5) | 2.3 ± 2.0 | 0.3 ± 0.2 | 0.1 ± 0.1 | 0.8 ± 0.9 | 6.4 ± 6.4 | 10.0 ± 0.8 |
| | | 1.3 | 0.3 | 0.03 | 0.2 | 2.8 | 4.6 |
| | | 0.6 – 5.5 | 0.2 – 0.7 | 0.02 – 0.3 | 0.1 – 2.2 | 2.2 - 17.2 | 4.4 - 22.0 |
| | Western (n = 4) | 12.3 ± 19.3 | 4.7 ± 5.1 | 0.1 ± 0.2 | 2.7 ± 2.3 | 57.9 ± 56.2 | 77.7 ± 57.6 |
| | | 3.5 | 4.0 | 0.08 | 2.3 | 39 | 67.0 |
| | | 1.9 – 41.3 | 0.1 - 10.8 | 0.03 – 0.4 | 0.3 - 5.2 | 15.8 – 137.2 | 19.1 – 157.3 |
| **Gizzard Shad** | Eastern (n = 5) | 2.4 ± 1.7 | 0.1 ± 0.1 | 0.04 ± 0.03 | 0.1 ± 0.1 | 2.1 ± 2.1 | 4.8 ± 3.2 |
| | | 2.2 | 0.2 | 0.02 | 0.2 | 1.3 | 4.4 |
| | | 0.4 – 4.6 | 0.03 - 0.2 | 0.01 - 0.9 | 0.09 – 0.2 | 0.7 - 5.7 | 1.3 – 9.8 |
| | East Central (n = 5) | 1.8 ± 1.5 | 0.2 ± 0.1 | 0.02 ± 0.01 | 0.4 ± 0.3 | 0.8 ± 0.4 | 3.2 ± 2.2 |
| | | 1.3 | 0.1 | 0.01 | 0.3 | 0.8 | 2.9 |
| | | 0.4 – 3.5 | 0.1 - 0.4 | 0.01 – 0.04 | 0.1 - 0.9 | 0.4 - 1.4 | 1.1 - 5.8 |
| | West Central (n = 5) | 4.0 ± 7.3 | 0.2 ± 0.1 | 0.01 ± 0.01 | 0.3 ± 0.2 | 0.9 ± 0.5 | 5.5 ± 7.9 |
| | | 0.6 | 0.3 | 0.01 | 0.2 | 0.8 | 2.1 |
| | | 0.6 - 17.1 | 0.06 - 0.4 | 0 – 0.02 | 0.2 - 0.6 | 0.5 - 1.8 | 1.1 – 19.7 |
| | Western (n = 8) | 17.0 ± 26.6 | 0.2 ± 0.3 | 0.05 ± 0.06 | 0.6 ± 0.5 | 3.1 ± 4.2 | 21.0 ± 30.8 |
| | | 4.5 | 0.1 | 0.03 | 0.8 | 1.7 | 5.9 |
| | | 0.2 - 51.7 | 0.02 – 0.8 | 0 - 0.2 | 0.2 - 1.7 | 0.3 - 12.9 | 0.7 - 66.9 |
| **Yellow Perch** | Eastern (n = 5) | 6.4 ± 3.6 | 0.08 ± 0.04 | 0.003 ± 0.002 | 0.5 ± 0.3 | 1.6 ± 1.0 | 8.5 ± 4.4 |
| | | 5.8 | 0.06 | 0.002 | 0.5 | 1.7 | 6.9 |
| | | 1.7 - 11.1 | 0.05 - 0.15 | 0.002 -0.005 | 0.2 - 0.8 | 0.4 - 2.8 | 4.2 - 14.9 |
| | East Central (n = 5) | 2.9 ± 1.7 | 0.5 ± 0.4 | 1.4 ± 1.2 | 1.1 ± 0.8 | 6.5 ± 6.2 | 12.4 ± 8.7 |
| | | 2.9 | 0.3 | 1.1 | 1.1 | 3.9 | 10.1 |
| | | 0.9 – 4.6 | 0.03 – 1.2 | 0.2 – 3.3 | 0.6 - 2.3 | 1.1 - 17.1 | 3.4 - 25.6 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | West Central (n = 5) | 9.9 ± 9.0 | 0.5 ± 1.0 | 0.4 ± 0.6 | 3.3 ± 3.1 | 12.0 ± 18.3 | 26.1 ± 27.6 |
| | | 8.3 | 0.2 | 0.2 | 3.6 | 2.6 | 13.5 |
| | | 0.8 – 24.9 | 0.03 - 2.1 | 0.07 - 1.4 | 0.08 - 8.0 | 0.8 – 43.7 | 8.5 – 74.9 |
| | Western (n = 7) | 22.2 ± 26.7 | 0.2 ± 0.1 | 0.2 ± 0.3 | 0.8 ± 0.5 | 2.1 ± 1.3 | 25.5 ± 27.4 |
| | | 9.5 | 0.1 | 0.08 | 0.5 | 2.0 | 12.1 |
| | | 2.7 – 41.5 | 0.04 - 0.5 | 0.001 - 0.8 | 0.4 – 1.5 | 0.4 – 3.6 | 5.7 – 46.1 |
| **Walleye** | Eastern (n = 5) | 6.2 ± 1.7 | 0.3 ± 0.2 | 0.4 ± 0.4 | 0.4 ± 0.3 | 2.8 ± 1.2 | 10.0 ± 2.3 |
| | | 5.9 | 0.1 | 0.3 | 0.3 | 2.6 | 10.5 |
| | | 4.0 – 8.0 | 0.07 - 0.6 | 0 – 0.6 | 0.1 - 0.9 | 1.6 - 2.2 | 7.3 – 11.0 |
| | East Central (n = 5) | 17.4 ± 0.9 | 0.1 ± 0.1 | 0.009 ± 0.004 | 0.2 ± 0.1 | 1.8 ± 0.5 | 19.6 ± 9.3 |
| | | 18.2 | 0.09 | 0.01 | 0.2 | 1.5 | 25.1 |
| | | 2.9 – 24.1 | 0.04 - 0.3 | 0.002 - 0.01 | 0.08 - 0.3 | 1.5 – 2.5 | 4.8 – 26.1 |
| | West Central (n = 3) | 37.7 ± 23.2 | 0.2 ± 0.06 | 0.07 ± 0.06 | 0.3 ± 0.02 | 4.9 ± 2.8 | 43.2 ± 26.9 |
| | | 49.3 | 0.2 | 0.02 | 0.3 | 5.1 | 51.9 |
| | | 4.8 – 59.0 | 0.1 - 0.2 | 0.01 - 0.06 | 0.3 - 0.3 | 2.1 - 7.6 | 13.0 - 64.6 |
| | Western (n = 10) | 3.3 ± 3.6 | 0.2 ± 0.1 | 0.06 ± 0.09 | 1.2 ± 0.7 | 2.6 ± 1.4 | 6.6 ± 4.6 |
| | | 2.3 | 0.1 | 0.01 | 1.2 | 2.5 | 7.0 |
| | | 0.7 – 13.3 | 0.06 – 0.5 | 0.004 - 0.3 | 0.2 - 2.2 | 0.5 - 5.3 | 3.7 – 16.4 |

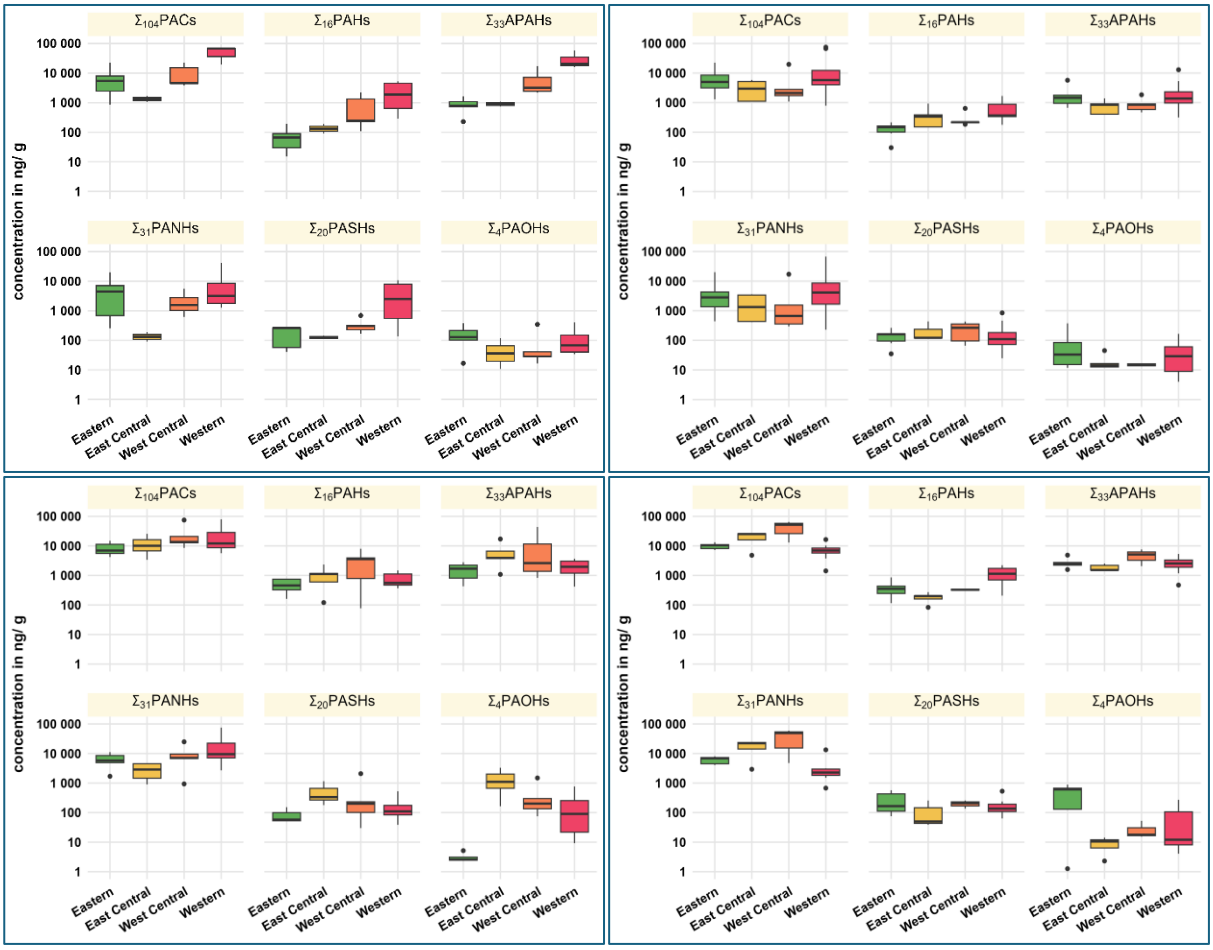



**Figure 5.1:** Log scaled boxplots of concentrations of major PAC classes (ng/ g, lipid-normalized) in four fish species from Lake Erie: (a) Freshwater Drum (top left), (b) Gizzard Shad (top right), (c) Yellow Perch (bottom left), and (d) Walleye (bottom right). Data are grouped by location (Eastern, East Central, West Central, and Western).

### 5.4.2 Spatial distribution of PACs in Lake Erie biota

Table 5.2 provides a statistical summary of the spatial variation in the concentrations of PACs in fish across the four regions of the lake. One-way ANOVA or the Kruskal-Wallis test was first used to determine if there were significant differences in mean concentration among regions. If a difference was found, a post-hoc test (Tukey's or Dunn's) was used to identify which specific regions had statistically different PAC concentrations.

**Table 5.2:** One way ANOVA/ Kruskal-Walli's and Post Hoc test results indicating the region-dependent concentration variations of PAC classes1.

| Species | Compound class | Test | F value/ chi-squared | P value | Region pair differences | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | Eastern | East Central | West Central | Western |
| **Freshwater Drum** | PACs | ANOVA + Tukey | 8.8 | 0.002 | a | a | a | b |
| | PAHs | | 6.5 | 0.007 | a | ab | ab | b |
| | APAHs | | 19.2 | 0.000 | a | ab | b | c |
| | PANHs | | 3.2 | 0.064 | a | a | a | a |
| | PASHs | | 4.0 | 0.035 | a | ab | ab | b |
| | PAOHs | | 0.7 | 0.561 | a | a | a | a |
| **Yellow Perch** | PACs | | 1.4 | 0.277 | a | a | a | a |
| | PAHs | | 1.0 | 0.395 | a | a | a | a |
| | APAHs | | 1.9 | 0.172 | a | a | a | a |
| | PANHs | | 2.8 | 0.073 | a | a | a | a |
| | PASHs | | 2.9 | 0.066 | a | a | a | a |
| | PAOHs | | 13.4 | 0.000 | b | a | ac | c |
| **Gizzard Shad** | PACs | | 1.0 | 0.407 | a | a | a | a |
| | PAHs | | 3.9 | 0.024 | a | ab | ab | b |
| | APAHs | | 1.4 | 0.286 | a | a | a | a |
| | PANHs | | 0.8 | 0.514 | a | a | a | a |
| | PASHs | | 0.5 | 0.660 | a | a | a | a |
| | PAOHs | | 0.7 | 0.550 | a | a | a | a |
| **Walleye** | PACs | Kruskal-Wallis + Dunn | 10.7 | 0.014 | ab | ab | a | b |
| | PAHs | ANOVA + Tukey | 10.4 | 0.000 | a | a | ab | b |
| | APAHs | | 1.6 | 0.220 | a | a | a | a |
| | PANHs | | 9.3 | 0.001 | ab | a | a | b |
| | PASHs | | 1.7 | 0.200 | a | a | a | a |
| | PAOHs | | 0.8 | 0.514 | a | a | a | a |

[1] Values with non-identical letters are significantly different from one another ($p<0.05$)

The mean concentrations of $\sum_{104}$PACs in freshwater drum from the western basin were significantly higher than those measured in the other three basins, which did not differ. For $\sum_{33}$APAHs, mean concentrations in freshwater drum from the eastern and east-central basins were statistically similar, yet both were significantly lower than those from the west-central and western basins. $\sum_{16}$PAH concentrations were also highest in freshwater drum collected from the western basin. With respect to $\sum_{20}$PASHs, mean concentrations in freshwater drum from the eastern and

east-central basins did not differ, whereas fish from the west-central and western basins showed significantly different concentrations from the eastern basin. Finally, no statistically significant variation was observed in $\sum_{31}$PANH or $\sum_{4}$PAOH concentrations among any of the basins for this fish. Taken together, these results suggest basin-dependent PAC exposure to freshwater drum. The movement of freshwater drum between basins is not well studied, but they are much more abundant in the western basin and likely do not move significantly across the lake.[41]

In yellow perch, the concentrations of all PACs except $\sum_{31}$PANHs were greatest in the two regions of the central basin. With the exception of $\sum_{4}$PAOHs, there were no significant differences in the concentrations of other PAC subclasses in the four regions of the lake. Yellow perch do not move significantly between basins,[40] so concentrations of these compounds are likely to be similar in the pelagic environment of the four regions of the lake compared to the benthic environment. According to the ANOVA results, in gizzard shad, all PAC groups except $\sum_{16}$PAHs remained statistically similar in all the regions, which was also evident in mean concentration profiles in Figure 5.1. Gizzard shad are mobile and migrate between basins depending on the season, for example, moving to the deeper waters of the eastern basin in the winter,[42] which likely explains the lack of concentration differences between basins. $\sum_{16}$PAHs concentrations were statistically different in eastern and western basin samples, while samples from both regions of the central basin were statistically similar to the other two basins.

Other than $\sum_{31}$PANHs, the concentrations of other PAC subclasses in walleye did not differ significantly among basins, similar to the gizzard shad, likely reflecting a highly mobile species that moves between basins.[43] It is known that walleye in the eastern basin do not migrate to the west and this is reflected in the significant concentration differences in $\sum_{31}$PANHs.[43] Unlike in

other species, where the $\sum_{104}$PAC mean concentration was highest in the western basin, in walleye it was highest in the central basin (western region).

The univariate analyses described above provided useful insights into differences in PAC concentrations and profiles among the four basins of Lake Erie. However, to obtain a more comprehensive understanding of spatial variability, two complementary statistical approaches, multivariate and cluster analyses, were employed. These methods enabled the identification of spatial trends and associations among the regions that were not evident from univariate comparisons alone. In particular, HCA grouped fish according to their PAC profiles, highlighting similarities and differences in exposure patterns across regions. Our group has previously applied these approaches to investigate the sources of PACs in Lake Ontario sediments.[37]

For the multivariate analysis, analytes were selected based on two criteria. First, only compounds with detection frequencies >75% were included, yielding 38 out of 50 HPACs. To reduce species-dependent variability, a one-way ANOVA test was performed across samples from the four regions. Analytes showing non-significant differences (p-value < 0.05) in at least two regions (Table III 7) were used as the original variables for the PCA and UMAP.

PCA results indicate that PC1, PC2, and PC3 explain 26%, 13% and 12% of the total variance, respectively, accounting for only 51% of the cumulative variance. Loading scores for the first five principal components are provided in Table III 7, and the corresponding biplots are shown in Figure 5.2. Positive PC1 is driven by PANHs such as quinoline isomers, carbazole, 3-methylisoquinoline and 2-methylacridine and PASHs like dibenzothiophene and 2,3,4-trimethylbenzothiophene meanwhile, negative PC1 is driven by halogenated HPACs (5-chloroindole and 2,3-dibromobenzothiophene) and carbazole derivatives such as 3,6-

dimethylcarbazole, 11H-benzo[c]carbazole and 9-methylcarbazole. PC2 is primarily influenced by PASHs, with the negative side influenced by HMW PANHs (11H-benzo[a]carbazole and dibenz[c,h]acridine). Positive PC3 is also driven by HMW PANHs, but negative PC3 is strongly associated with alkylated PASHs.

The basin-dependent effect is most pronounced along PC1, with a clear separation of western and eastern basin samples from those of the central basin, particularly for yellow perch and gizzard shad. Gizzard shad from the central basin are enriched in halogenated HPACs, while eastern and western basin gizzard shad are more influenced by PANHs and alkylated PASHs. In freshwater drum, PC1 values are lower for western and west-central samples with high influence from halogenated HPACs compared to eastern and east-central samples that were driven by unsubstituted and alkylated PANHs and PASHs. This likely reflects benthic habitat variations along the lake's depth profile and differences in point sources of halogenated and non-halogenated HPACs. Walleye samples exhibit the greatest basin-independent differences, reflective of the high level of movement between basin for this species.[43] However, compared to other samples, western basin walleye samples have high positive PC1 and PC3 values, indicating those samples are enriched in PANHs and alkylated PASHs. PC2 and PC3 appear to capture other ecological variables, such as trophic position and habitat differences; for example, walleye and gizzard shad overlap along PC2, consistent with predator–prey transfer of halogenated HPACs, while walleye and yellow perch overlap along PC3, indicating shared habitat exposures and similar accumulation patterns of halogenated vs alkylated HPACs. Interestingly, the elevated concentrations of HMW PANHs in the east central yellow perch samples have high PC3 scores clustering away from other samples, suggesting a possible unique exposure event.

Unlike PCA, which captures global variance, UMAP emphasizes local data structure based on the given hyperparameters, making it more sensitive to basin-driven clusters.[37] Therefore, in comparison, UMAP analysis produced a more distinct clustering of HPAC profiles among species across the three basins. With respect to freshwater drum, yellow perch, and gizzard shad, there is a clear separation between central basin samples and the other two basins that cluster with each other. The basin-independent clustering of walleye from four regions is also more evident in UMAP outputs (Figure 5.2). Overall results of multivariate analysis indicate that HPACs profiles of most species are mostly dominated by the basin-dependent variations, while trophic position and mobility can have a greater impact on some species.

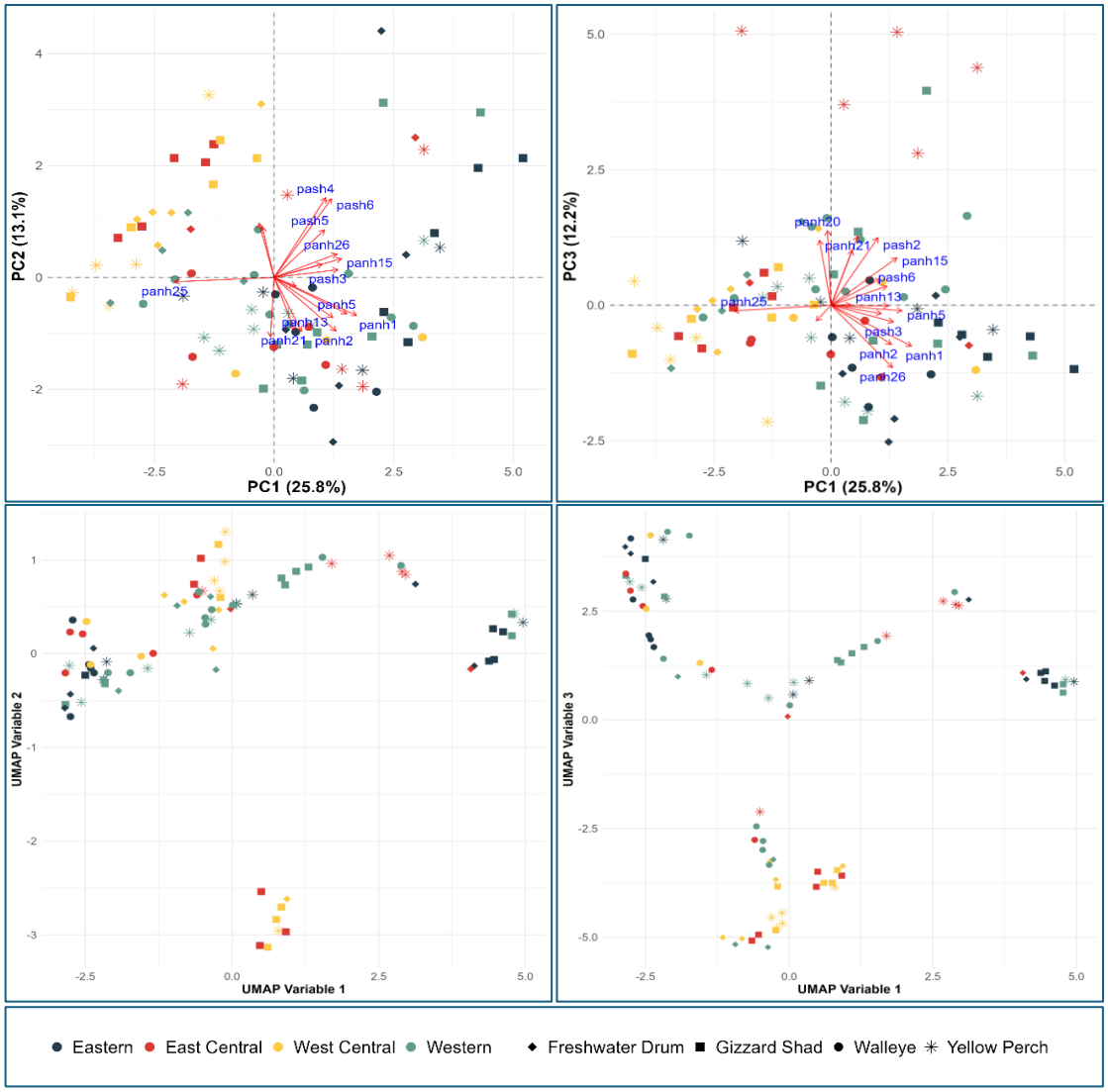



**Figure 5.2:** Multivariate analysis of PAC profiles using principal component analysis (PCA) and uniform manifold approximation and projection (UMAP). PCA score plots are shown for (a) PC1 vs. PC2 (top left) and (b) PC1 vs. PC3 (top right), while UMAP projections are shown for (c) variable 1 vs. variable 2 (bottom left) and (d) variable 1 vs. variable 3 (bottom right). Symbols represent fish species, and colours indicate sampling regions (Eastern, East Central, West Central, and Western Lake Erie). (panh1: quinoline, panh11: acridine, panh12: 9-methylcarbazole, panh13: carbazole, panh15: 2-methylacridine, panh16: 3,6-dimethylcarbazole, panh2: isoquinoline, panh20: 11H-benzo[a]carbazole, panh21: dibenz[c,h]acridine, panh25: 5-chloroindole, panh26: 3-bromoisoquinoline, panh5: 3-methylisoquinoline, pash18:2,3-dibromobenzothiophene, pash2: 2,3-dimethylbenzothiophene, pash3: 2,3,4-trimethylbenzothiophene, pash4: dibenzothiophene, pash5: 4-methyldibenzothiophene, pash6: 2,4-dimethyldibenzothiophene).

Although multivariate analysis highlighted important underlying HPAC profiles in the Lake Erie system, it was limited to only eleven of the 50 analyzed HPACs after the preselection criteria used. To provide a more comprehensive assessment of the impact of a higher number of HPACs, HCA was performed using two approaches. The first approach considered all species simultaneously across different locations (Figure III 6), which helped identify species-independent variations within basins and tended to yield more stable clusters due to the larger sample size. However, species-specific metabolism and differences in bioaccumulation could obscure basin-specific contamination patterns. To address this and as a second approach, separate HCAs were conducted for each species (Figure 5.3).

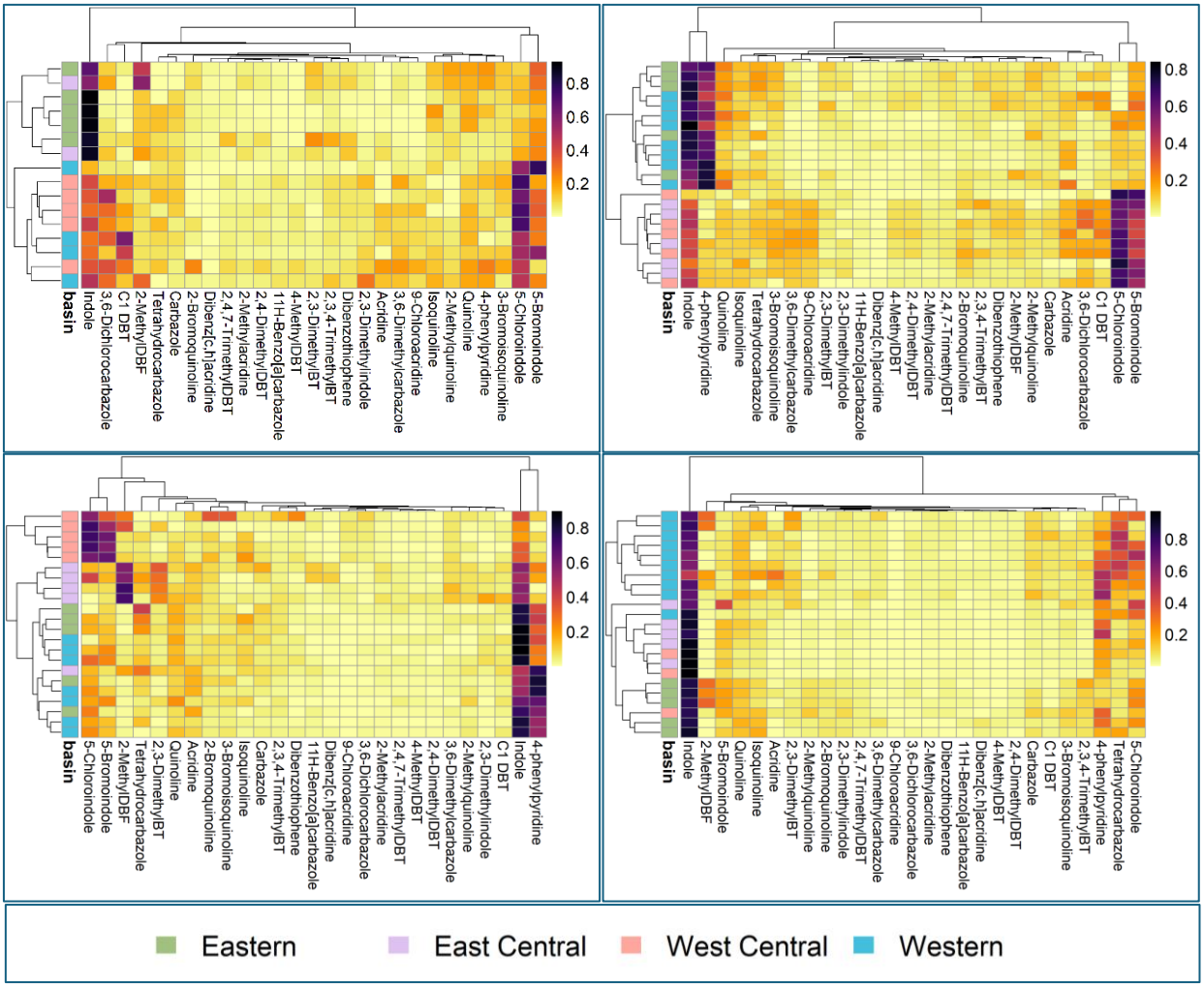



**Figure 5.3:** Heatmaps generated from hierarchical cluster analysis (HCA) results of (a) Freshwater Drum (top left), (b) Gizzard Shad (top right), (c) Yellow Perch (bottom left), (d) Walleye (bottom right).

Based on the HPAC profiles, freshwater drum formed two main clusters: the eastern/east-central cluster and a western/west-central cluster. Notably, there was a distinct separation between unsubstituted PANH and halogenated PANH across these clusters. In eastern/east-central samples, unsubstituted PANHs such as indole, quinoline, and carbazole predominated, whereas chlorinated and brominated PANHs were more abundant in western/west-central samples. A similar trend was observed for dibenzothiophenes, with unsubstituted forms dominating the eastern cluster and

halogenated forms in the western. As a benthic omnivore,[32] the freshwater drum data suggest that the western basin benthic environment is more influenced by industrial and agricultural inputs. The elevated levels of unsubstituted PANHs in the eastern basin may reflect bacterial degradation and other weathering processes acting on halogenated PANHs as water flows easterly.

The HPAC profiles of gizzard shad and yellow perch exhibited trends distinct from those of freshwater drum. These samples also formed two clusters, but in this case, the clusters corresponded to the western/eastern basin and the central basin. Unsubstituted HPACs, such as indole and 4-phenylpyridine, were higher in the western/eastern cluster, whereas halogenated HPACs predominated in the central cluster. These patterns suggest that HPAC profiles differ between pelagic and benthic environments in Lake Erie, and that these two pelagic species are exposed to similar HPAC compositions due to their ecology and trophic positions. Notably, the basin-specific HPAC patterns observed in freshwater drum, yellow perch, and gizzard shad were consistent across both multivariate analysis based on seven selected HPACs and cluster analysis using 28 analytes.

Consistent with the multivariate analysis, hierarchical cluster analysis (HCA) revealed that walleye exhibited the greatest variation in HPAC concentration profiles between basins. Concentrations of indole and its halogenated derivatives had a particularly strong influence on the clustering compared to most other analytes. Samples from the east-central and west-central basins were interspersed with those from the other basins, whereas the western and eastern basins displayed distinct pollution profiles. These patterns likely reflect long-range migration and the mixing of pelagic and benthic food sources, which influence HPAC exposure in walleye.

The primary sources of HPACs in Lake Erie likely include water coming from the upper Great Lakes, industrial discharges, agricultural runoff, legacy contamination sites, atmospheric deposition, and petroleum, automotive, and industrial waste.[4,44] Overall, the western basin exhibited the highest contamination levels of parent HPACs among the three basins. This pattern can be attributed to major inflows such as the Detroit and Maumee Rivers, which transport PACs from highly industrialized/urbanized and agricultural regions, including the greater Detroit metropolitan area, as well as agricultural runoff and wastewater effluent from the Ohio area.[45] In the pelagic environment, unsubstituted parent HPACs dominate in the eastern basin relative to halogenated and alkylated HPACs. Sources such as industrial effluents and historical waste sites (e.g., Love Canal) likely contribute to these eastern basin concentrations of PACs.[46]

A potential important source of PACs to the central basin of Lake Erie is the Cleveland Harbor and the adjacent Cuyahoga River inflow in Ohio.[45] This region is known to deliver petroleum-rich industrial waste, oil refinery effluents, and legacy contamination from steel mills into the lake.[47] This explains the high concentrations of more nonpolar, persistent HPACs, such as halogenated PANHs, observed in the benthic freshwater drum in this study. The central basin also receives high inputs of agricultural runoff, which likely contributes to the elevated abundance of halogenated HPACs in the water column compared with the eastern and western basins.

Hierarchical cluster analysis across all species indicates that the central basin, particularly its western region, is enriched in brominated HPACs. Brominated HPAC concentrations were lower in the top predator, suggests a reduced potential for bioaccumulation. However, this finding may also indicate the presence of other brominated organic compounds with the potential to bioaccumulate.

Relatively high levels of 4-phenylpyridine in the pelagic zones of the western and eastern basins point to contamination by low-molecular-weight HPACs, likely originating from industrial sources. This compound was not detected in freshwater drum and walleye and is consistent with its high-water solubility and low bioaccumulation potential.

Among basin-species combinations, yellow perch from the central basin exhibited elevated concentrations of alkylated HPACs, including 2-methyldibenzofuran, 2,3-dimethylbenzothiophene, and 3,6-dimethylcarbazole. This pattern suggests greater exposure of yellow perch to alkylated PACs and supports a petrogenic source of these compounds. In contrast, comparatively lower concentrations of these HPACs (except indole) in walleye relative to other species indicate limited bioaccumulation, consistent with expected reduced trophic transfer.

## 5.5 CONCLUSIONS

This study showed that HPACs were the most dominant type of PACs in fish from Lake Erie. Out of the three types of HPACs, PANHs were much abundant in all fish, confirming their high-water solubility and bioavailability compared to PASHs and PAOHs. Halogenated HPACs and alkylated dibenzothiophenes were in higher concentrations in freshwater drum samples compared to other species, indicating their abundance in the benthic environment. Basin-specific differences were evident, with the western basin consistently exhibiting the highest contaminant loads, largely driven by inflows from the Detroit and Maumee Rivers. Multivariate analysis demonstrated distinct basin-dependent contaminant signatures. Hierarchical cluster analysis further explored the variations in PAC profiles in the different basins. Collectively, the results highlight the importance of integrating HPACs into environmental monitoring programs while also providing evidence of basin- and species-specific concentration differences. The chemometric approaches used in our

study were paramount in revealing patterns of contamination in biota from different regions of the lake. In addition, this study establishes new baseline data to support ecological risk assessments in the Great Lakes. While dissolved-phase concentrations in water were not measured in this study because fish tissues analyzed here were archived samples originally from a previous study, future work should incorporate water measurements to better resolve exposure pathways and link abiotic sources with bioaccumulation patterns. Such efforts should also utilize bioindicator species representing both benthic and pelagic food web pathways, integrate sediment, air and water quality parameters to more effectively link sources to exposure, and employ temporal monitoring to assess trends in HPAC contamination. Such integrated approaches will improve our ability to evaluate ecological and human health risks and to guide management actions in the Great Lakes.

## 5.6 ACKNOWLEDGEMENTS

This work was supported by a Natural Sciences and Engineering Research Council Discovery Grant to GTT (Grant number: RGPIN-2024-04947) and ATF (RGPIN-2020-05500) and a University of Manitoba Graduate Fellowship to NNV.

## 5.7 DECLARATION OF GENERATIVE AI AND AI ASSISTED TECHNOLOGIES IN THE MANUSCRIPT PREPARATION PROCESS

During the preparation of this work the authors used OpenAI. (2026). ChatGPT and Google. (2026). Gemini in order to generate and proofread the R codes. After using these tools/services the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

## 5.8 REFERENCES


(1) Marvin, C. H.; Berthiaume, A.; Burniston, D. A.; Chibwe, L.; Dove, A.; Evans, M.; Hewitt, L. M.; Hodson, P. V.; Muir, D. C. G.; Parrott, J.; Thomas, P. J.; Tomy, G. T. Polycyclic Aromatic Compounds in the Canadian Environment: Aquatic and Terrestrial Environments. *Environmental Pollution*. Elsevier Ltd September 15, **2021**. https://doi.org/10.1016/j.envpol.2021.117442.

(2) Xia, Z.; Idowu, I.; Marvin, C.; Thomas, P. J.; Johnson, W.; Francisco, O.; Stetefeld, J.; Crimmins, B.; Fry, M.; Tomy, G. T. Identification of Halogenated Polycyclic Aromatic Hydrocarbons in Biological Samples from Alberta Oil-Sands Region. *Chemosphere* **2019**, 215, 206–213. https://doi.org/10.1016/j.chemosphere.2018.10.050.

(3) Lammel, G. Polycyclic Aromatic Compounds in the Atmosphere – A Review Identifying Research Needs. *Polycycl. Aromat. Compd.* **2015**, 35 (2–4), 316–329. https://doi.org/10.1080/10406638.2014.931870.

(4) Berthiaume, A.; Galarneau, E.; Marson, G. Polycyclic Aromatic Compounds (PACs) in the Canadian Environment: Sources and Emissions. *Environmental Pollution*. Elsevier Ltd January 15, **2021**. https://doi.org/10.1016/j.envpol.2020.116008.

(5) Wallace, S. J.; de Solla, S. R.; Head, J. A.; Hodson, P. V.; Parrott, J. L.; Thomas, P. J.; Berthiaume, A.; Langlois, V. S. Polycyclic Aromatic Compounds (PACs) in the Canadian Environment: Exposure and Effects on Wildlife. *Environmental Pollution*. Elsevier Ltd October 1, **2020**. https://doi.org/10.1016/j.envpol.2020.114863.

(6) Stout, S. A.; Emsbo-Mattingly, S. D.; Douglas, G. S.; Uhler, A. D.; McCarthy, K. J. Beyond 16 Priority Pollutant PAHs: A Review of PACs Used in Environmental Forensic Chemistry. *Polycycl. Aromat. Compd.* **2015**, 35 (2–4), 285–315. https://doi.org/10.1080/10406638.2014.891144.

(7) Andersson, J. T.; Achten, C. Time to Say Goodbye to the 16 EPA PAHs? Toward an Up-to-Date Use of PACs for Environmental Purposes. *Polycycl. Aromat. Compd.* **2015**, 35 (2–4), 330–354. https://doi.org/10.1080/10406638.2014.991042.

(8) Ghosh, P.; Mukherji, S. Fate, Detection Technologies and Toxicity of Heterocyclic PAHs in the Aquatic and Soil Environments. *Science of the Total Environment.* Elsevier B.V. September 20, **2023**. https://doi.org/10.1016/j.scitotenv.2023.164499.

(9) Machado, M. E.; Nascimento, M. M.; Bomfim Bahia, P. V.; Martinez, S. T.; Bittencourt de Andrade, J. Analytical Advances and Challenges for the Determination of Heterocyclic Aromatic Compounds (NSO-HET) in Sediment: A Review. *TrAC - Trends in Analytical Chemistry.* Elsevier B.V. **2022**. https://doi.org/10.1016/j.trac.2022.116586.

(10) Landis, M. S.; Studabaker, W. B.; Patrick Pancras, J.; Graney, J. R.; Puckett, K.; White, E. M.; Edgerton, E. S. Source Apportionment of an Epiphytic Lichen Biomonitor to Elucidate the Sources and Spatial Distribution of Polycyclic Aromatic Hydrocarbons in the Athabasca Oil Sands

Region, Alberta, Canada. *Science of the Total Environment* **2019**, 654, 1241–1257. https://doi.org/10.1016/j.scitotenv.2018.11.131.

(11) Zou, Y.; Wang, L.; Christensen, E. R. Problems in the Fingerprints Based Polycyclic Aromatic Hydrocarbons Source Apportionment Analysis and a Practical Solution. *Environmental Pollution* **2015**, 205, 394–402. https://doi.org/10.1016/j.envpol.2015.05.029.

(12) Steinman, A. D.; Cardinale, B. J.; Munns, W. R.; Ogdahl, M. E.; Allan, J. D.; Angadi, T.; Bartlett, S.; Brauman, K.; Byappanahalli, M.; Doss, M.; Dupont, D.; Johns, A.; Kashian, D.; Lupi, F.; McIntyre, P.; Miller, T.; Moore, M.; Muenich, R. L.; Poudel, R.; Price, J.; Provencher, B.; Rea, A.; Read, J.; Renzetti, S.; Sohngen, B.; Washburn, E. Ecosystem Services in the Great Lakes. *Journal of Great Lakes Research.* International Association of Great Lakes Research June 1, **2017**, pp 161–168. https://doi.org/10.1016/j.jglr.2017.02.004.

(13) Gewurtz, S. B.; Diamond, M. L. Distribution and Burdens of Bioaccumulative Contaminants in the Lake Erie Food Web: A Review. *Environmental Reviews.* **2003**, pp 141–160. https://doi.org/10.1139/a03-014.

(14) Allinger, L. E.; Reavie, E. D. The Ecological History of Lake Erie as Recorded by the Phytoplankton Community. *Journal of Great Lakes Research.* September **2013**, pp 365–382. https://doi.org/10.1016/j.jglr.2013.06.014.

(15) Gewurtz, S. B.; Backus, S. M.; Bhavsar, S. P.; McGoldrick, D. J.; De Solla, S. R.; Murphy, E. W. Contaminant Biomonitoring Programs in the Great Lakes Region: Review of Approaches and Critical Factors. *Environmental Reviews* **2011**, 19 (1), 162–187. https://doi.org/10.1139/A11-005.

(16) Baumann, P. C.; Whittle, D. M. The Status of Selected Organics in the Laurentian Great Lakes: An Overview of DDT, PCBs, Dioxins, Furans, and Aromatic Hydrocarbons; **1988**; Vol. 11.

(17) Zhang, X.; Bowman, D. T.; Diamond, M. L.; Helm, P.; Jobst, K. J.; Hao, C.; Kleywegt, S.; Zhang, Z. F.; Marvin, C.; Zhang, X. Contribution of Coal Tar Sources to Polycyclic Aromatic Compounds and Associated Ecological Risk in Lake Ontario Sediments: Inference from a Novel Marker. *Environ. Sci. Technol.* **2025**. https://doi.org/10.1021/acs.est.4c10360.

(18) Li, A.; Zhou, S.; He, H.; Guo, J.; Rockne, K. J.; Sturchio, N. C.; Giesy, J. P. Polyhalogenated Carbazoles in Sediments of Lower Laurentian Great Lakes and Regional Perspectives. ACS *ES and T Water* **2022**, 2 (9), 1544–1554. https://doi.org/10.1021/acsestwater.2c00191.

(19) Guo, J.; Li, Z.; Ranasinghe, P.; Bonina, S.; Hosseini, S.; Corcoran, M. B.; Smalley, C.; Rockne, K. J.; Sturchio, N. C.; Giesy, J. P.; Li, A. Spatial and Temporal Trends of Polyhalogenated Carbazoles in Sediments of Upper Great Lakes: Insights into Their Origin. *Environ. Sci. Technol*. **2017**, 51 (1), 89–97. https://doi.org/10.1021/acs.est.6b06128.

(20) Heuvel, C. E.; Zhao, Y. M.; Fisk, A. T. Food Web Structure across Basins in Lake Erie, a Large Freshwater Ecosystem. *Canadian Journal of Fisheries and Aquatic Sciences* **2025**, 82. https://doi.org/10.1139/cjfas-2024-0028.

(21) Vitharana, N. N.; Halldorson, T.; Xia, Z.; Francisco, O.; Marvin, C.; Thomas, P. J.; Liebzeit, E.; Lucas, A. M.; Moradi, V.; Tomy, G. T. A Validated Approach for Analysis of Heterocyclic Aromatic Compounds in Sediment Samples. *J. Chromatogr. A* **2024**, 1718. https://doi.org/10.1016/j.chroma.2024.464723.

(22) Idowu, I.; Francisco, O.; Thomas, P. J.; Johnson, W.; Marvin, C.; Stetefeld, J.; Tomy, G. T. Validation of a Simultaneous Method for Determining Polycyclic Aromatic Compounds and Alkylated Isomers in Biota. *Rapid Communications in Mass Spectrometry* **2018**, 32 (3), 277–287. https://doi.org/10.1002/rcm.8035.

(23) The Fitness for Purpose of Analytical Methods : A Laboratory Guide to Method Validation and Related Topics, Second edition, **2014**.; Magnusson, B., Örnemark, U., Eds.

(24) Vitharana, N. N. PCA_UMAP_LE. GitHub. https://github.com/NipuniV/PCA_UMAP_LE.git (accessed 2026-01-07).

(25) Vitharana, N. N. HCA_LE. GitHub. https://github.com/NipuniV/HCA_LE.git (accessed 2026-01-07).

(26) Han, T. H.; Lee, J. H.; Cho, M. H.; Wood, T. K.; Lee, J. Environmental Factors Affecting Indole Production in *Escherichia Coli*. *Res. Microbiol.* **2011**, 162 (2), 108–116. https://doi.org/10.1016/j.resmic.2010.11.005.

(27) Sullivan, D. W.; Gad, S. E. Indole. In Encyclopedia of Toxicology: Third Edition; Elsevier, **2014**; pp 1030–1031. https://doi.org/10.1016/B978-0-12-386454-3.00860-5.

(28) Ma, Q.; Zhang, X.; Qu, Y. Biodegradation and Biotransformation of Indole: Advances and Perspectives. *Frontiers in Microbiology.* Frontiers Media S.A. November 1, **2018**. https://doi.org/10.3389/fmicb.2018.02625.

(29) Vrzalová, A.; Vrzal, R.; Nádvorník, P.; Šebela, M.; Dvořák, Z. Modulation of Aryl Hydrocarbon Receptor Activity by Halogenated Indoles. *Bioorg. Med. Chem*. **2024**, 114. https://doi.org/10.1016/j.bmc.2024.117964.

(30) Roseman, E. F.; Mills, E. L.; Forney, J. L.; Rudstam, L. G. Evaluation of Competition between Age-0 Yellow Perch (Perca Flavescens) and Gizzard Shad (Dorosoma Cepedianum) in Oneida Lake, New York 1; **1996**.

(31) Jake, M.; Zanden, V.; Cabana, G.; Rasmussen, J. B. Comparing Trophic Position of Freshwater Fish Calculated Using Stable Nitrogen Isotope Ratios (d15N) and Literature Dietary Data; **1997**.

(32) Heuvel, C. E.; Haffner, G. D.; Zhao, Y.; Colborne, S. F.; Despenic, A.; Fisk, A. T. The Influence of Body Size and Season on the Feeding Ecology of Three Freshwater Fishes with Different Diets in Lake Erie. *J. Great Lakes Res*. **2019**, 45 (4), 795–804. https://doi.org/10.1016/j.jglr.2019.05.001.

(33) Habib, I.; Singha, K.; Hossain, M. Recent Progress on Pyridine N-Oxide in Organic Transformations: A Review. *ChemistrySelect*. John Wiley and Sons Inc February 3, **2023**. https://doi.org/10.1002/slct.202204099.

(34) Lemos, A.; Lemaire, C.; Luxen, A. Progress in Difluoroalkylation of Organic Substrates by Visible Light Photoredox Catalysis. *Advanced Synthesis and Catalysis*. Wiley-VCH Verlag April 1, **2019**, pp 1500–1537. https://doi.org/10.1002/adsc.201801121.

(35) Wu, Y.; Tan, H.; Zhou, C.; Crimmins, B. S.; Holsen, T. M.; Chen, D. Bioaccumulation and Spatiotemporal Trends of Polyhalogenated Carbazoles in Great Lakes Fish from 2004 to 2016. *Environ. Sci. Technol.* **2018**, 52 (8), 4536–4545. https://doi.org/10.1021/acs.est.8b00427.

(36) Dreij, K.; Lundin, L.; Le Bihanic, F.; Lundstedt, S. Polycyclic Aromatic Compounds in Urban Soils of Stockholm City: Occurrence, Sources and Human Health Risk Assessment. *Environ. Res*. **2020**, 182. https://doi.org/10.1016/j.envres.2019.108989.

(37) Vitharana, N. N.; Halldorson, T.; Dereviankin, M.; Paddock, B.; Gerring, A.; Marvin, C.; Helm, P.; Thomas, P. J.; Idowu, I. G.; Tomy, G. T. Heterocyclic Aromatic Compounds in Lake Ontario Sediments: Spatiotemporal Distribution and New Approaches to Fingerprinting. *Science of the Total Environment* **2025**, 986. https://doi.org/10.1016/j.scitotenv.2025.179802.

(38) Li, M.; Larter, S. R. Potential Bias in the Isolation of Pyridinic Nitrogen Fractions from Crude Oils and Rock Extracts Using Acid Extraction and Liquid Chromatography; **2001**. https://doi.org/https://doi.org/10.1016/S0146-6380(01)00062-6.

(39) Metcalfe, C. D.; Metcalfe, T. L.; Riddle, G.; Douglas Haffner, G. Aromatic Hydrocarbons in Biota from the Detroit River and Western Lake Erie. *J. Great Lakes Res*. **1997**, 23 (2), 160–168. https://doi.org/10.1016/S0380-1330(97)70893-1.

(40) Gewurtz, S. B.; Lazar, R.; Haffner, G. D. Comparison of Polycyclic Aromatic Hydrocarbon and Polychlorinated Biphenyl Dynamics in Benthic Invertebrates of Lake Erie, USA. *Environ. Toxicol. Chem.* **2000**, 19 (12), 2943–2950. https://doi.org/10.1002/etc.5620191215.

(41) Bur, M. T. Growth, Reproduction, Mortality, Distribution, and Biomass of Freshwater Drum in Lake Erie. *J. Great Lakes Res*. **1984**, 10 (1), 48–58. https://doi.org/10.1016/S0380-1330(84)71806-5.

(42) Slagle, Z. Report of the Lake Erie Forage Task Group Members; **2020**.

(43) Matley, J. K.; Faust, M. D.; Raby, G. D.; Zhao, Y.; Robinson, J.; MacDougall, T.; Hayden, T. A.; Fisk, A. T.; Vandergoot, C. S.; Krueger, C. C. Seasonal Habitat-Use Differences among Lake

Erie's Walleye Stocks. *J. Great Lakes Res*. **2020**, 46 (3), 609–621. https://doi.org/10.1016/j.jglr.2020.03.014.

(44) Li, W.; Park, R.; Alexandrou, N.; Dryfhout-Clark, H.; Brice, K.; Hung, H. Multi-Year Analyses Reveal Different Trends, Sources, and Implications for Source-Related Human Health Risks of Atmospheric Polycyclic Aromatic Hydrocarbons in the Canadian Great Lakes Basin. *Environ. Sci. Technol.* **2021**, 55 (4), 2254–2264. https://doi.org/10.1021/acs.est.0c07079.

(45) Marvin, C. H.; Painter, S.; Charlton, M. N.; Fox, M. E.; Thiessen, P. A. L. Trends in Spatial and Temporal Levels of Persistent Organic Pollutants in Lake Erie Sediments. *Chemosphere* **2004**, 54 (1), 33–40. https://doi.org/10.1016/S0045-6535(03)00660-X.

# CHAPTER 6 : HABITAT AND DIET AFFECT BIOACCUMULATION OF HOMOCYCLIC AND HETEROCYCLIC AROMATIC COMPOUNDS IN LAKE ERIE FISH

This chapter builds on my earlier work and presents the bioaccumulation of PACs in ten species of fish from the western basin of Lake Erie. I was also able to discern the overarching factors driving bioaccumulation of PACs in these species. Because of the large data set, I used a few advanced chemometric tools to help identify PAC profiles in the different fish species. Redundancy analysis (RDA) was conducted to assess the impact of environmental and biological parameters on the bioaccumulation of PACs, and, to my knowledge, this is the first study to use RDA for this purpose. Trophic magnification and biomagnification factors for each PAC were also calculated, and for most PACs, these values were reported for the first time using field data.

A version of this chapter is published as Vitharana Nipuni N, Halldorson Thor, Paddock Brett, Tomy Thane, Marvin Chris, Fisk A. T., Heuvel C. E., Tomy Gregg T., Habitat and Diet Affect Bioaccumulation of Homocyclic and Heterocyclic Aromatic Compounds in Lake Erie Fish, Journal of Hazardous Materials, 2026, 142044. Copyright © 2026. B.V. Elsevier. Reprinted with permission.

## 6.1 ABSTRACT

Polycyclic aromatic compounds (PACs) are a broad class of structurally diverse chemicals that originate from both natural and anthropogenic sources. A comprehensive suite of PACs consisting of 49 homocyclic PACs and 55 heterocyclic PACs (HPACs) was analysed in nine fish species from the western basin of Lake Erie to identify their distribution and bioaccumulation patterns. Lipid-corrected total PAC (∑PACs) concentrations ranged from 6 to 110 µg/g among species. Alkylated polycyclic aromatic hydrocarbons and polycyclic aromatic nitrogen heterocycles were the dominant subclasses, accounting for over 60% of the $\sum_{104}$PACs in eight out of the nine species. Unconstrained ordination approaches revealed distinct HPAC contamination profiles among fish species, largely influenced by habitat and feeding ecology. In contrast, homocyclic PACs did not show species-specific clustering, suggesting they reflect lake wide rather than species-driven contamination patterns. Redundancy analysis (RDA) identified fish length and habitat use, assessed using $\delta^{13}C$ and $\delta^{34}S$, as significant drivers of HPACs profiles, whereas trophic position, determined from $\delta^{15}N$, had little influence. More than 75% of analytes had non-significant TMFs, with significant TMFs exhibiting low $R^2$ values, indicating low biomagnification potential and even biodilution. Notably, RDA explained only one-third of the total variance, indicating that unexplained biological and environmental factors contribute to HPACs variability. Overall, results suggest HPAC accumulation in freshwater ecosystems is driven by environmental exposure pathways rather than dietary transfer.

## 6.2 ENVIRONMENTAL IMPLICATIONS

Simultaneously analyzing a broad suite of polycyclic aromatic compounds (PACs), revealed distinct bioaccumulation patterns between homocyclic and heterocyclic PACs in Lake Erie fish. Our findings suggest sediment contamination, habitat- and diet-driven exposure pathways, and point sources of specific PACs. The elevated levels of alkylated PAHs and nitrogen-containing PACs, relative to the 16 priority PAHs, highlight the need to include these compounds in routine monitoring and risk assessment programs. By doing so, this will improve our understanding of PAC sources, their environmental fate, and their ecological effects across freshwater food webs.

## 6.3 INTRODUCTION

As the world's largest freshwater ecosystem, the Laurentian Great Lakes (*hereafter* Great Lakes) have been subjected to significant and ongoing contaminant inputs.[1–3] The shallowest of the five Great Lakes, Lake Erie has a large and important commercial and recreational fisheries and high recreational use.[4,5] Bathymetrically and limnologically, the lake is divided into three distinct regions: western, central, and eastern basins, with depth increasing from west to east.[6] The western basin experiences substantial nutrient inputs from agricultural runoff, leading to eutrophication and anoxic zones in the central basin of the lake, and also contamination inputs from industry.[7] The ecological and biological diversity of fish populations within the western basin, therefore, provides a relevant and sensitive ecosystem for investigating the bioaccumulation and fate of organic chemicals.[6]

Polycyclic Aromatic Compounds (PACs) are one of the major classes of organic pollutants in the aquatic environment.[8] This family of compounds can be divided into several subclasses: polycyclic aromatic hydrocarbons (PAHs), alkylated polycyclic aromatic hydrocarbons (APAHs), halogenated polycyclic aromatic hydrocarbons (HPAHs) and hetero-polycyclic aromatic compounds (HPACs).[9] Both the United States Environmental Protection Agency (US-EPA) and the Canadian Environmental Protection Act 1999 identify 16 PAHs as priority pollutants, and hence, there have been a number of studies done to identify the presence of PAHs in Lake Erie fish.[10] Also, there have been several studies that focus on determining the bioaccumulation and biomagnification potential of those 16 PAHs in different aquatic food webs.[11,12] Other than PAHs and APAHs, very little attention has been given to understanding the environmental fate and behavior of other subclasses of PACs in biota samples.

Hetero-polycyclic aromatic compounds (HPACs) are an understudied subclass of PACs. Like other types of PACs, HPACs originate from pyrogenic and petrogenic sources and the limited toxicological data suggests that these compounds may have similar or higher toxicity relative to PAHs.[13–15] Therefore, it is important to explore the persistence, bioaccumulative and toxic (PB&T) properties of these chemicals in different environmental systems.

Homocyclic and heterocyclic PACs have been identified in different abiotic environmental matrices such as sediments, snow and air samples.[16–18] Also, a few studies have reported concentrations of a select number of HPACs in Great Lakes biota. [19–21] However, the overall understanding of the bioaccumulation potential of PACs, including those containing nitrogen, sulfur, and oxygen, remains limited. Previously, our group reported on the spatial differences in HPACs among fish from different basins of Lake Erie, showing that HPAC concentrations varied with the ecological and biological characteristics of the species.[22]

To further our understanding of the fate and behavior of PACs, we examined the bioaccumulation potential of HPACs in nine fish species from the western basin of Lake Erie that feed on different habitats and trophic positions. The species included emerald shiner (*Notropis atherinoides*), spottail shiner (*Notropis hudsonius*), gizzard shad (*Dorosoma cepedianum*), yellow perch (*Perca flavescens*), white bass (*Morone chrysops*), walleye (*Sander vitreus*), round goby (*Neogobius melanostomus*), white sucker (*Catostomus commersonii*) and freshwater drum (*Aplodinotus grunniens*). To assess levels in the water column as a baseline for the ecosystem, HPACs were analyzed in Dreissenid mussels (most likely *Dreissena bugensis*), a filter feeder of the pelagic zone.[6] In total, 55 HPACs, 16 PAHs and 33 APAHs were measured across all samples along with $\delta^{13}C$ and $\delta^{34}S$, which provide insights on habitat use, and $\delta^{15}N$, a measure of trophic position. Two

unconstrained ordination approaches, principal component analysis (PCA) and uniform manifold approximation and projection (UMAP) analysis, and one constrained ordination approach, redundancy analysis (RDA), were used to identify natural clustering of fish based on HPACs profiles. Additionally, lipid-normalized biomagnification factors (BMF) for specific prey-predator relationships and trophic magnification factors (TMF) were calculated for selected HPACs to further elucidate the bioaccumulation potential of HPACs in the food web.

## 6.4 MATERIALS AND METHODS

### 6.4.1 Chemicals

All polycyclic aromatic analytes and deuterated PAHs (internal standards) were purchased from Chiron (AS Stiklestadveien 1, 7041 Trondheim, Norway) and Sigma Aldrich Canada Ltd (Oakville, Ontario, Canada). High-purity (Optima grade) organic solvents, silica gel (923 grade, 100 - 200 mesh), alumina (60 - 325 mesh), Ottawa sand, and anhydrous sodium sulfate, diatomaceous earth (DE) were purchased from Fisher Chemicals (Ottawa, Canada). All the target analytes and internal standards used in this study are summarized in Table IV 1 with their reference codes and CAS numbers listed in Table IV 2.

### 6.4.2 Biota samples collection and stable isotope analysis

Fish samples were collected from different locations in the western basin of Lake Erie between May and October 2019 as part of the Cooperative Science Monitoring Program (Forage Task Group 2023).[6] In total 61 fish samples were analysed for PACs that included emerald shiner (n = 7), spottail shiner (n = 8), gizzard shad (n = 8), yellow perch (n = 8), white bass (n = 7), walleye (n = 10), round goby (n = 4), white sucker (n = 5) and freshwater drum (n = 4). Dreissenid mussels

(n=5) were collected biweekly from May to October 2019, at six sites (two sites/basin) during limnological surveys of Lake Erie by the Ontario Ministry of Natural Resources and Forestry (see Heuvel *et al*. 2023 for site locations and details);[23] because of their small sample size multiple mussels were combined to make a single sample. The details of length, weight, and exact locations were available for most of the fish (Table VI.3), and the detailed procedures followed during sample collection, storage, stable isotope (δ13C, δ15N, and δ34S) analysis, and trophic position (Table IV 3) calculations were published in Heuvel *et al*.[6] The sample collection and stable isotope analysis procedures are also described in Appendix IV.

### 6.4.3 Extraction, cleanup and gas chromatography tandem mass spectrometry analysis

Sample extraction for PAC analysis was carried out using accelerated solvent extraction using 1:1 mixture of DCM and ethyl acetate as extraction solvents. Samples were then subjected to several cleanup techniques prior to gas chromatography tandem mass spectrometry (GC-MS/MS) analysis. Details on the extraction and cleanup procedures and GC-MS/MS operating conditions are provided in our previous work.[24,25] A detailed description of the methods and quality control procedures is also included in Appendix IV. Method validation was carried out for all the analytes according to the Eurachem Guide -The Fitness for the Purpose of Analytical Methods.[26] The method validation procedure and performance characteristics data for homocyclic PACs can be found in Idowu *et al*.[24] and the method validation data for HPACs (Table IV 4) can be found in Vitharana *et al*.[27] and Appendix IV. The average lipid data for each species can be found in Table VI.5.

Our laboratory complies with ISO 17027:2018 requirements by participating in proficiency testing exercises twice per year. In the most recent Wepal-Quasimeme (Netherlands) proficiency test, z-

scores for 33 PACs, including PAHs, APAHs and dibenzothiophenes, quantified in a biota certified reference material, were consistently below a z-score value of 2. This implies that for those analytes, our method is suitable for its intended purpose. To our knowledge, no existing proficiency testing program captures more than these 33 PACs. However, our own in-house QC measures are designed to ensure good data quality objectives.

### 6.4.4 Multivariate analysis

The statistical analysis was performed using RStudio 2025.09.1+401 "Cucumberleaf Sunflower" released on 2025-09-23) for Windows. The raw data and codes used for all the R analysis can be found on GitHub.[28,29] Principal component analysis and UMAP were performed to identify natural clustering among species based on HPAC profiles. Then, to relate HPACs distributions to stable isotope patterns ($\delta^{13}C$, $\delta^{15}N$ and $\delta^{34}S$) and the size of the specimen (wet weight and total length) RDA was performed as a constrained ordination analysis. All three multivariate analysis were performed using (a) only HPACs, (b) only PAHs and APAHs and (c) considering all 104 PACs to identify the differences between the profiles driven by homocyclic and heterocyclic PACs.

The sample size for the different species ranged from 4 to 8. A post hoc simulation-based power analysis was conducted using repeated subsampling of individuals within species. Results showed that detection power exceeded 94% at $n = 3$ per species and reached 100% at $n \geq 4$, indicating that our sampling intensity is more than sufficient for robust RDA inference.

For multivariate analysis, the dataset was first filtered based on compound detectability. To evaluate the influence of detection-frequency filtering on multivariate structure, ordinations generated using different analyte detection percentage thresholds (70%, 75%, 80%, and 90%) were compared using Procrustes analysis. High Procrustes correlations ($r = 0.69$ - $0.86$, $p = 0.001$)

indicated that overall sample relationships were largely preserved across these four detection percentages. Therefore, 80% detection percentage was selected to filter out analytes prior to statistical tests because it provided an optimal balance between ordination stability and retention of chemically informative analytes. Concentrations below the LOD were replaced with a value of 0.01 ng/g to remove the zero induced inflation during multivariant analysis. Lipid normalized concentration data were Hellinger transformed. Subsequently, relationships within the dataset were explored using PCA and UMAP.

For the constrained ordination, the explanatory environmental parameters ($\delta^{13}C$, $\delta^{15}N$, $\delta^{34}S$ weight and length) were standardized and z-scores were calculated. To verify that the data met the RDA prerequisite of linearity between response (HPACs concentrations) and explanatory matrices (environmental parameters), detrended correspondence analysis was performed on the Hellinger transformed data. A resulting gradient length of $< 3$ supported the suitability of the linear model. Multicollinearity among the explanatory variables was ruled out using the variance inflation factor (VIF $< 10$) analysis. Statistical significance was determined via permutation tests (999 times) to identify the statistical significance i) in overall analysis, ii) by RDA axes, and iii) by explanatory variables. The final results were visualized using an RDA biplot of RDA1 and RDA2.

### 6.4.5 Quantification of biomagnification

To determine the BMF and TMF of each PAC the following equations were used.

$$BMF_{TP} = ([Predator]_{Lipid\ Corrected} / [Prey]_{Lipid\ Corrected}) / (TP_{predator} / TP_{prey}) \quad \text{Equation 1}$$

$$\log C_{Lipid\ corrected} = a + b \times TP \quad \text{Equation 2}$$

$$TMF = 10^{b} \quad \text{Equation 3}$$

where, TP is the trophic position of each fish which was calculated in our previous publication from $\delta^{15}N$ using a two source model that incorporated benthic and pelagic baselines using $\delta^{13}C$, Heuvel *et al*., 2025.[6] $C_{lipid\ corrected}$ is lipid corrected concentration in the predator and prey species, and a and *b* are intercept and the slope of the linear relationship between $C_{lipid\ corrected}$ and TP across all species. BMF and TMF > 1 indicate an increase in concentrations between predator and prey and for the entire food web, respectively.[30]

## 6.5. RESULTS AND DISCUSSION

### 6.5.1 Profiles and concentrations of polycyclic aromatic compounds in biota

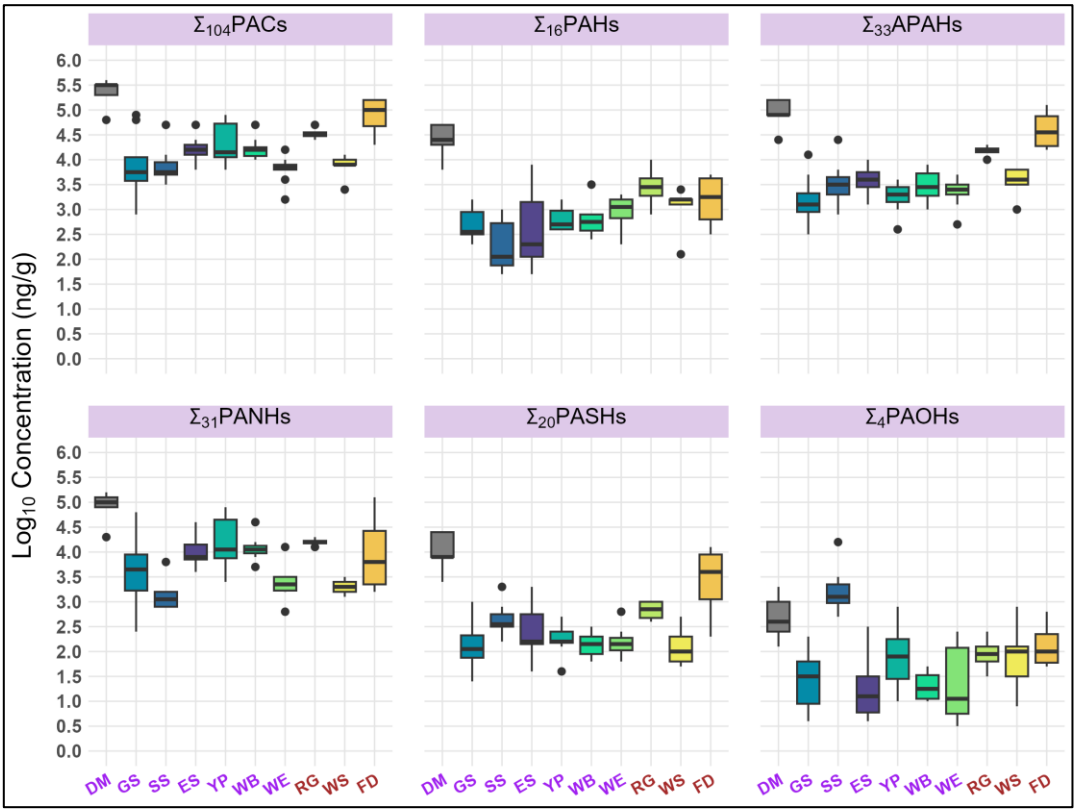



**Figure 6.1:** Box and whisker plots of log-transformed concentrations (lipid-corrected) of different PAC types in all the species of interest (pelagic species: purple labels and benthic species: brown labels) in order of increasing trophic positions from the western basin of Lake Erie. Species are labeled as dreissenid mussels (DM), gizzard shad (GS), spottail shiner (SS), emerald shiner (ES), yellow perch (YP), white bass (WB), walleye (WE), round goby (RG), whitesucker (WS) and freshwater drum (FD).

The concentrations of PACs in fish from Lake Erie are provided in Figure 6.1 and Table IV 6. In general, the concentrations of PACs varied by species and compound class. In most species, the total body burden of PACs was dominated by PANHs and APAHs; the US-EPA 16 priority PAHs

accounted for less than 5% of the total PAC burden across all ten species. For clarity, PAC concentrations in benthic and pelagic species will be discussed in separate sections.

(a) Benthic food web

The rank order of the arithmetic mean $\sum_{104}$PACs in benthic species was freshwater drum > round goby ≈ white sucker (see Table SI6). Apart from freshwater drum, PANHs were the most abundant PAC groups in the benthic species. Concentrations of $\sum_{31}$PANHs accounted for *ca*. 49% of $\sum_{104}$PACs in round goby (17 µg/g, lw) and *ca*. 34 % in white sucker (2 µg / g, lw). However, in freshwater drum (9 µg/g, lw), the concentrations of $\sum_{31}$PANHs were less than 10% of ∑PACs. The most abundant PANH in all species was indole (which accounted for ~12% of $\sum_{104}$PACs in white sucker but less than 2% in other fish), and unlike in white sucker, the other two species had elevated concentrations of halogenated PANHs (41% in round goby and 8% in freshwater drum of $\sum_{104}$PACs). Low molecular weight (LMW) PANHs such as 4-phenylpyridine, quinoline isomers and carbazole were also found in detectable quantities in all species but made up a small percentage (< 6%) of $\sum_{104}$PACs.

Compared to PANHs, the $\sum_{20}$PASHs (5 to 0.09 µg/g, lw) and $\sum_{4}$PAOHs (0.1 to 0.09 µg/g, lw) were detected in much lower concentrations in the three benthic fish. For these species, the PASH profiles were dominated by alkylated dibenzothiophenes in freshwater drum and halogenated PASHs in round goby, accounting for 88% and 72% $\sum_{20}$PASHs, respectively. However, in white sucker there was not an overwhelmingly dominant PASH. Dibenzothiophene was the most abundant PAOH in all three fish, accounting for *ca*. 50% of the $\sum_{4}$PAOHs.

Lipid-adjusted concentrations of $\sum_{16}$PAHs and $\sum_{33}$APAHs ranged from 3 to 1 and 40 to 4 µg/g (lw), respectively, in benthic species. Freshwater drum had the highest concentration of $\sum_{16}$PAHs

and $\sum_{33}$APAHs while white sucker had the lowest concentrations. In freshwater drum (39 µg / g, lw), $\sum_{33}$APAHs accounted for *ca*. 50 % of the $\sum_{104}$PACs. Alkylated fluorenes, pyrenes and chrysene were the dominant homocyclic compounds in freshwater drum and occurred at higher concentrations than in other benthic species.

Benthic fish generally showed higher $\sum_{16}$PAH and $\sum_{33}$APAH concentrations than pelagic fish (*see below*), consistent with the greater hydrophobicity of PAHs and APAHs relative to their heterocyclic counterparts. This stronger affinity for sediment-associated organic matter, combined with the direct contact and feeding habits of benthic species, results in greater exposure to sediment-bound PACs compared to pelagic species.

(b) Pelagic food web

In the pelagic food web, except for the PAOHs, the highest concentration of all PAC subclasses were detected in Dreissenid mussels. These mussels are filter feeders and are commonly used as a trophic level 2 baseline of the pelagic food web, indicating the low biomagnification potential of these compounds. The PANHs (median: 98 µg/g, lw) accounted for 34% of the $\sum_{104}$PACs in this species. Unlike other species, there were detectable amounts of all 20 PASHs in mussels, and the 16 PAHs were *ca.* x10 higher in Dreissenid mussels (median: 23 µg/g, lw) compared to all fish species. Unlike in most fish, ∑31PANHs concentrations were lower than ∑33APAHs in Dreissenid mussels. These higher concentrations and greater range of detectable compounds is due to lower biotransformation capacity of organic contaminants, such as PCBs, in Dreissenid mussels compared to fish.[31] As well, due to their high octanol-water partition (Kow) coefficients (>3), PACs preferentially adsorb to suspended particles and sediments and helps to explain the observed elevated concentrations in filter feeders such as mussels.

When only the pelagic fish are considered, the $\sum_{104}$PACs concentrations tended to increase with trophic position, except for the top predator, walleye (see Figure 6.1 and Table IV 6). Among the pelagic fish, emerald shiner and gizzard shad had the highest and lowest median concentration of $\sum_{104}$PACs (16 and 6 µg/g, lw), respectively.

Except for spottail shiner and walleye, $\sum_{104}$PAC concentrations in all other fish were largely driven by $\sum_{31}$PANHs, which accounted for more than 70% of the total burden in yellow perch, gizzard shad and white bass. In contrast, $\sum_{31}$PANHs contributed only about 20% to the $\sum_{104}$PAC concentration in spottail shiner. In all species, indole was the most abundant N-based compound, accounting for 10 to 60% of $\sum_{104}$PACs. Across all pelagic species except the spottail shiner, 4-phenylpyridine and quinoline isomers (which are LMW PANHs) were present at relatively higher concentrations than the other analytes. The presence of N increases the polarity of PANHs relative to PAHs and APAHs, leading to greater water solubility and bioavailability. The high-water solubility is consistent with the high levels of the LMW PACs in fish as the primary route of exposure for these compounds is via direct uptake through gills and skin[32].

In contrast to PANHs, the median concentration of $\sum_{20}$PASHs remained relatively consistent among pelagic fish (0.4 to 0.1 µg/g, lw), with the highest concentration detected in emerald shiner. The median concentration of $\sum_{4}$PAOHs was the lowest across all PAC groups (1 to 0.01 µg/g, lw). Due to the high percentage of non-detects, the % RSTD of $\sum_{4}$PAOHs exceeded 100% in all species. However, an interesting finding was a trophic dilution of PAOHs with respect to both pelagic and benthic species.

Across the six pelagic fish species, the median $\sum_{16}$PAHs and $\sum_{33}$APAHs concentrations ranged from 1 to 0.1 and 62 to 6 µg / g (lw), respectively. Phenanthrene was the only PAH that was

detected in all the species. C2-chrysene was the dominant homocyclic PAH detected across all pelagic species, with concentrations ranging from 1 to 0.3 µg/g, lw (2 to 23% of $\sum_{104}$PACs). The spottail shiner was the only pelagic fish where the concentration of $\sum_{33}$APAHs exceeded that of $\sum_{31}$PANHs.

Despite differences in trophic position, the concentrations of $\sum_{16}$PAHs, $\sum_{33}$APAHs and $\sum_{20}$PASHs remained relatively constant across all fish species. This lack of difference suggests that APAHs and PASHs behave similarly to PAHs in biological systems, such as efficient metabolic breakdown by organisms, low dietary assimilation, and physicochemical properties that favor elimination rather than accumulation.

To assess the potential for adverse effects from the measured PACs exposure, we compared our fish tissue concentrations (ranging from 0.1 to 5 ug/g, $\sum_{16}$PAH) against established developmental toxicological endpoints. The comparison included concentrations known to cause developmental toxicity in rainbow trout (*Oncorhynchus mykiss*, $\sum_{37}$PAH and APAH = 0.7 µg/g, ww),[33] low hatching success and ventricular ballooning in pacific herring (*Clupea pallasi*, $\sum_{42}$PACs = 0.03 to 3 µg/g)[34,35] and microscopic hepatobiliary changes in red snapper (*Lutjanus campechanus*, $\sum_{46}$PAH and APAH and dibenzothiophene 0.2 – 8.5 µg/g).[36] Overall, the concentrations measured in our study are well within those known to cause adverse effects. We estimated a site-specific hazard quotient (HQ) for freshwater drum by comparing our measured $\sum_{16}$PAH and $\sum_{33}$APAH tissue concentrations (*ca.* 2.9 µg/g, ww) against the established developmental toxicity threshold of 0.7 µg/g. The resulting HQ of 4.2 indicates a potential for adverse developmental effects. It should be emphasized, however, because PAHs and APAHs comprise *ca.* 40% of the total measured PAC burden in freshwater drum, the true cumulative risk is likely to be significantly

higher. This is particularly important because studies have confirmed the higher toxic effect of PACs due to alkylation, halogenation, and presence of hetero atoms, although the number of studies that focus on the toxicity of these subclasses of compounds is much less than the homocyclic PACs.[37,38] Taken together, the high relative abundance of non-priority PACs (>95%) observed in this study underscores the need for ongoing monitoring and highlights the importance of including APAHs and HPACs in risk assessment frameworks to accurately evaluate their potential hazards.

### 6.5.2 Species and habitat driven clustering based on multivariate analysis

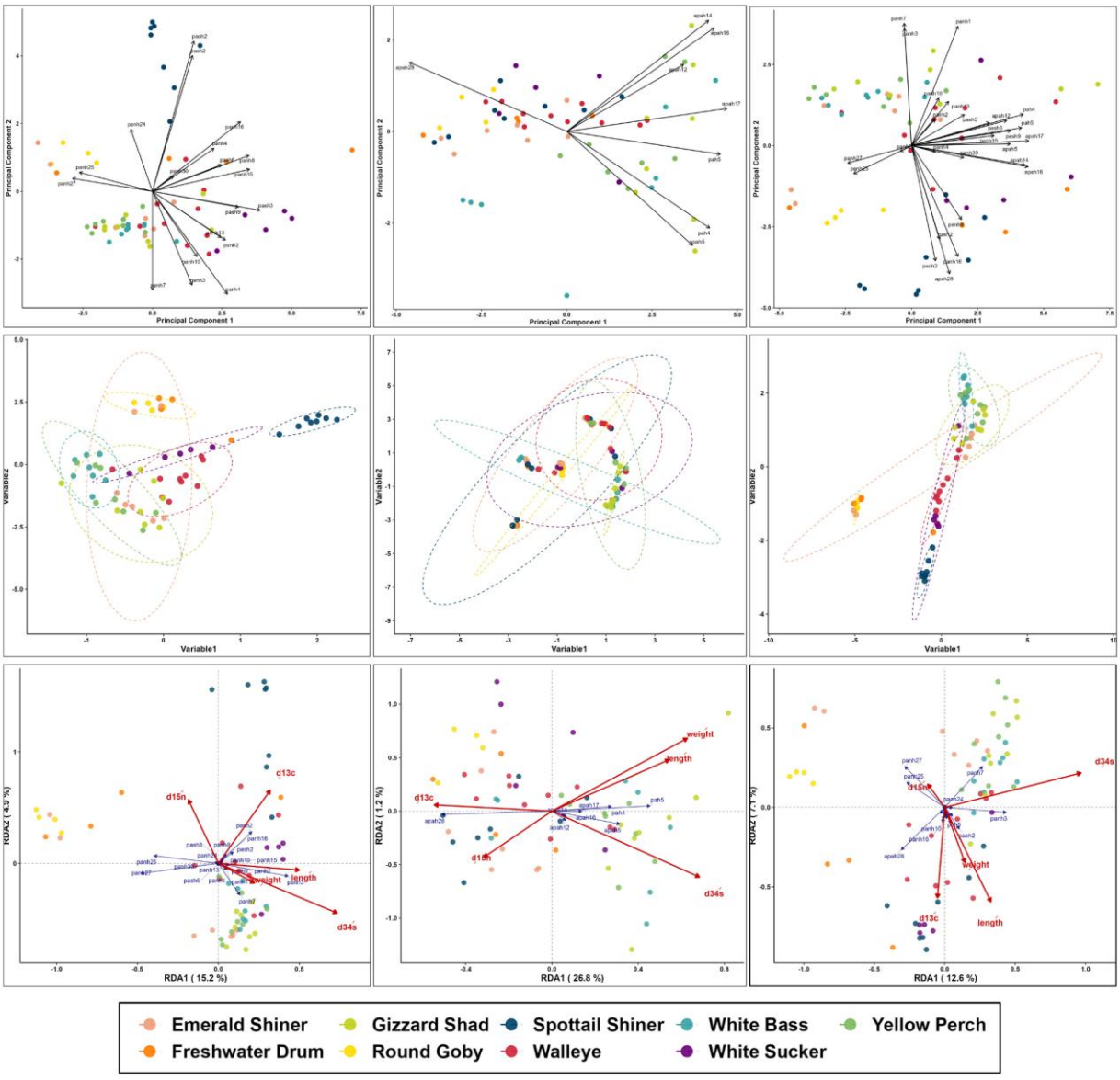



**Figure 6.2:** PCA biplots between (a) PC1 and PC2 (top panel), UMAP plots between (b) variable 1 and variable 2 (middle panel-horizontal), and RDA plot between (c) RDA1 and RDA2 (bottom panel). The left panels are for the HPACs-based analyses, the middle panels (vertical) are for homocyclic-based analyses, and the right panels are for total PAC-based analyses

Environmental contaminant concentrations in fish samples are governed by multiple factors, including chemical and physical properties, environmental parameters (e.g., pH, temperature, salinity), and species-specific variations in diet, distribution, biotransformation, and excretion.[39] To elucidate the distribution patterns of PACs concentrations across species, three multivariant analysis (PCA, UMAP and RDA) to three distinct compound groups: (a) HPACs, (b) PAHs and APAHs and (c) all PACs combined (Figure 6.2, Table IV 7). Specific details for each multivariate analysis are summarised in Table 6.1.

**Table 6.1:** Information on the multivariate analyses done for the identification of the species clustering based on the PAC profiles

| | **HPACs** | **PAHs and APAHs** | **All PACs** |
|---|---|---|---|
| Total number of analytes | 54 | 49 | 104 |
| No of original variables (analytes > 80% detection limit) | 19 | 8 | 27 |
| **Principal component analysis (PCA)** | | | |
| Cumulative variance for the first three principal components | 56% | 92% | 55% |
| High loading analytes PCA1 | indole, carbazole, 2,3-dimethylindole, 3,6-dimethylcarbazole, benzo(h)quinoline | 9-methylphenanthrene, C2-chrysene, phenanthrene, benzo(g,h,i)perylene, fluorene | 9-methylphenanthrene, 3-methylphenanthrene, 2-methylphenanthrene, fluorene, phenanthrene |
| High loading analytes PCA2 | 2-methyldibenzofuran, 2,3-dimethylbenzothiophene, quinoline, 4-phenylpyridine, indole | C4-naphthelene, 2-methylphenanthrene, 3-methylphenanthrene, Fluorene, C2-chrysene | C2-chrysene, 4-phenylpyridine, quinoline, indole, 2-methyldibenzofuran |
| **Uniform manifold approximation (UMAP)** | | | |
| Number of neighbours | 15 | 5 | 12 |

| | | | |
|---|---|---|---|
| Minimum distance | 0.05 | 0.05 | 0.01 |
| Number of components | 3 | 3 | 3 |
| Cluster overlap | Low | High | Moderate |
| **Redundancy analysis (RDA)** | | | |
| % variance explained by the explanatory variables | 25% | 30% | 25% |
| Model significance (p_value) | 0.001 | 0.001 | 0.001 |
| Significance of the explanatory variables (p_values) | | | |
| $\delta^{13}C$ | 0.008 | 0.041 | 0.024 |
| $\delta^{15}N$ | 0.079 | 0.310 | 0.196 |
| $\delta^{34}S$ | 0.001 | 0.003 | 0.001 |
| Length | 0.008 | 0.004 | 0.006 |
| weight | 0.071 | 0.086 | 0.085 |
| % cumulative variance for RDA1 and RDA2 | 80% | 94% | 78% |

The PCA results show that homocyclic PACs are more closely associated with each other compared with the HPACs as shown by their lower cumulative variance. Conversely, when HPACs were incorporated into the model, the variability increases suggesting that these compounds originate from multiple sources and exhibit a more complex environmental behaviour. In the HPAC-based analysis, PC1 is predominantly driven by PANHs which distinctly separate the three benthic fish species from each other, while the pelagic fish cluster more closely to each other. Interestingly, sediment dwelling white sucker shows the highest PC1 scores, whereas the round goby which primarily feeds on dreissenid mussels shows the lowest, consistent with the high abundance of PANHs in sediments.[18] The freshwater drum, which feeds on both these species[40] spans both ends of PC1. In contrast, PC2 is influenced by all three HPACs classes and reveals a

clear separation of spottail shiner from the other species. This pattern is mainly driven by 2-methyldibenzofuran and 2,3-dimethylbenzothiophene rather than PANHs implying the unique influence of PASHs and PAOHs in differentiating spottail shiner from other fish.

For the homocyclic PACs, PCA analysis indicate that both high molecular weight (HMW) pyrogenic PACs (Benzo(g,h,i)perylene and C2-chrysene) and low molecular weight (LMW) petrogenic PACs (9-methylphenanthrene, phenanthrene, and fluorene) contribute to the main principal components. Although the species dependent variation is much lower in this analysis than the HPACs analysis, there is evidence of zone-specific chemical distribution of fish. C2 chrysene is influencing the benthic organisms while LMW PAHs influence pelagic species. However, unlike other pelagic fish, emerald shiner and white bass have low PC1 scores indicating possible differences in metabolism of PAHs and APAHs. The top predator, walleye who travels between the two zones is spread across the benthic and pelagic species. These trends align with known chemical properties: the high log Kow values of HMW PACs drive their binding to sediments, while LMW PACs remain more volatile and water soluble.[41] PCA did not reveal any clear species separation.

Comparison of the two PCAs reveals two exposure scenarios: HPACs are linked to species-specific patterns, likely driven by biotransformation, and diverse environmental sources while PAHs and APAHs display a mixed petrogenic and pyrogenic origin likely derived from atmospheric and watershed inputs. The combined PAC-PCA successfully captures both these patterns. With this combined model, PC1 is interpreted as the overall PAC burden, and PC2 serves as a discriminator for the heterocyclic versus homocyclic balance. We acknowledge that certain individuals (e.g., white sucker, spottail shiner, and emerald shiner) present as outliers far from their clusters, which

may be attributed to site-specific contamination, sampling variability or individual biotransformation capacity differences.

Given that the PCA analysis for HPACs and total PACs both showed a moderate cumulative variance of ca. 50%, we employed UMAP analysis to further explore the local structure of the data. The goal was to determine if samples would cluster according to their similarities in ecological niches rather than just the linear combination of the variables. The three UMAP analyses were performed iteratively, with hyperparameter adjustments until the optimal species separation was achieved (Figure IV 1 shows the UMAP outputs between variables 1 and variable 3). As illustrated in Figure 6.2, UMAP confirmed the distinct species-specific HPAC contamination profiles.

When UMAP analysis carried out with only HPACs is considered, even with the highest number of neighbours (15) comparing the three UMAP analyses, gave the best species-specific clustering. Just as in PCA, the output shows tight clusters for benthic species like the white sucker, whose profile is attributed to its sediment-ingesting ecology. The round goby formed a distinct cluster separate from the white sucker, while the freshwater drum samples were spread between the two, indicating dietary exposure sourced from both benthic types. Conversely, the pelagic forage fish (emerald shiner, gizzard shad, yellow perch, and white bass) formed highly overlapping clusters, suggesting common diets and habitats define their HPAC exposure. The one exception among the forage fish is spottail shiners, whose HPAC profile clustered separately from all other pelagic fish. The wide spread of walleye across the UMAP confirms its opportunistic apex predator[42] and its HPAC profile is likely a direct reflection of the contamination profiles of its prey. Overall, these results imply that feeding behaviors produce distinct HPAC signatures.

Even with a minimal number of neighbors, UMAP analysis of the homocyclics showed high species overlap, indicating these compounds are uniformly distributed throughout the Lake Erie basin. This homogeneity is likely a result of their widespread sources, high abundance in the lake and consistent biotransformation across species. The combined PAC UMAP analysis showed moderate species overlap and successfully separated fish into two ecological zones. Our analyses strongly suggest that homocyclic PACs are best suited for monitoring temporal trend and basin-wide control, while HPACs are more reliable for identifying species- and habitat-specific areas of concern.

We then conducted an RDA to integrate the species-dependent PAC clustering with key ecological factors, including stable isotope signatures ($\delta^{13}C$, $\delta^{15}N$ and $\delta^{34}S$) and fish size (see Table IV 3). In general, the models explained 25-30% of the variance, which is considered acceptable for ecological data analysis.[43,44] However, a majority of the variance is unexplained and would be related to a combination of factors, including biological (e.g., biotransformation, microbial breakdown, age, sex), abiotic (e.g., temperature, turbidity), and chemical (e.g., photolysis, hydrolysis).

Based on permutation test on individual predictor variable, $\delta^{13}C$, $\delta^{34}S$ and body length were consistent predicators of PAC concentrations in fish. The relationships with $\delta^{13}C$ and $\delta^{34}S$ indicate that habitat is an important determinate of PAC concentrations, where benthic organisms had higher concentrations (*see below*). Body size is often a driver of contaminant concentrations in fish, with concentrations increasing with size,[45] although additional work is needed to understand how important this is for PACs. Interestingly, $\delta^{15}N$ was not a significant predictor of concentration indicating that PAC compounds do not biomagnify in fish, likely driven by high biotransformation.

Given that RDA1 and RDA2 captured over 75% of the constrained variance and were the only axes with a significant effect ($p<0.05$), our subsequent data interpretation will focus solely on these two constrained axes.

In the HPAC-based RDA, the strong influence of $\delta^{34}S$ on RDA1 suggests that this axis reflects differences between benthic and pelagic feeding pathways. This isotope has also been used to track sewage and other anthropogenic sources (Tucker *et al*. 1999),[46] although our lack of understanding of $\delta^{34}S$ sources and dynamics in the Great Lakes makes specific conclusions difficult. In addition to sulfur stable isotope signatures, variations in carbon sources also appear to influence RDA1. Fish with distinct $\delta^{34}S$ and $\delta^{13}C$ accumulated different HPAC mixtures.

Effect of variations in carbon source on HPACs profiles evident along the RDA2 axis. More positive $\delta^{13}C$ values align with benthic carbon sources, whereas more negative values reflect pelagic phytoplankton sources. The benthic species, white sucker, showed high positive RDA2 scores consistent with its feeding on benthic detritus. Conversely, pelagic species (emerald shiner, yellow perch, and gizzard shad) show lower RDA2 scores. The position of the white sucker is heavily influenced by petrogenic HPACs, specifically 2,3-dimethylbenzothiophene, carbazole (and its derivatives), and isoquinoline. This implies that these contaminants are key markers of the benthic food web.

Although a benthic species, round goby showed the most negative RDA2 scores. This is likely explained by their consumption of the filter-feeding dreissenid mussels,[47] which had the highest concentration of HPACs; none of the other fish species feed on these mussels. Since these mussels filter particulate organic matter, including phytoplankton, from the water column, the round goby exposure profile shifts resulting in a unique HPAC signature. Halogenated PANHs, such as 5-

bromoindole, 5-chloroindole, and 3-bromoisoquinoline are key contaminants associated with this mussel-based food web. The source of these HPACs is likely from industrial inputs or microbial processes. The RDA also visually identifies the freshwater drum's mixed diet signature as its position spreads across the chemical profiles of the round goby and white sucker.

An unexpected observation was the high RDA1 and RDA2 scores for the spottail shiner. Its high RDA1 scores is likely due to its near-shoreline habitat and consumption of benthic invertebrates and organic detritus resulting in high $\delta^{13}C$ values.[48] However, the species is notable separated along RDA2, potentially due to its unique $\delta^{34}S$ signature arising with microbial-mediated sulfur reduction near shallow shorelines. There were no unique analytes specifically observed with the spottail shiner cluster, but the novel patterns and concentrations in this fish warrants further study in small forage fish.

The comparison across all multivariate techniques reveals clear differences in HPACs behavior. As observed in the unconstrained multivariant analyses, homocyclic PACs consistently yielded the lowest separation in the RDA results, suggesting their distribution is more uniform and less influenced by habitat-specific factors. Conversely, for all PACs, the RDA provided more distinct species-specific and habitat-driven clusters; RDA1 was primarily governed by $\delta^{34}S$ and $\delta^{13}C$ mainly influenced RDA2; both of these stable isotopes are traces of carbon source and habitat. Consistent with PCA and UMAP, the RDA confirmed walleyes role as an opportunistic predator.[49] The broad spread of walleye samples along RDA1 reflects their variable diet of both benthic and pelagic fish, confirming that the walleye's HPAC profile is a direct reflection of its prey's contamination profiles. Overall, the RDA model can help delineate fish feeding ecology and underscores how contaminant profiles are highly sensitive to these ecological pathways.

### 6.5.3 Bioaccumulation of polycyclic aromatic compounds in Lake Erie food web

The RDA suggested that trophic position, based on $\delta^{15}N$, did not significantly influence HPAC distribution in Lake Erie fish. To assess this further, we calculated BMFs and TMFs for all 104 individual PACs (see Table IV 7 and Table IV 8). BMF values for most compounds (*ca*. 80%) were near zero for common predator-prey feeding relationships (e.g. walleye-perch and walleye-emerald shiner) implying that these compounds are being biodiluted between prey and consumer.

Interpretation of BMFs requires caution due to potential analytical artifacts. For example, concentrations of halogenated PANHs and HMW APAHs are near or below the limit of detection in white sucker thereby generated spuriously high BMFs; these were also confirmed to be non-biomagnifying using the TMF approach. With walleye as the predator, acridine and 3-methyl quinoline BMFs were orders of magnitude higher when round goby was the prey (3586 and 325, respectively) compared to spottail shiner (1 and 0). This discrepancy aligns with RDA and UMAP findings that the unique HPAC profiles of round goby and spottail shiner are driven by habitat and biological differences, not trophic transfer. Overall, the high proportion of BMFs near zero across feeding relationships suggests effective biotransformation and elimination of HPACs.

Although most PACs should have little species-specific biomagnification, significant TMFs values for a limited number of chemicals provide some evidence for biomagnification. Of the 104 PACs examined, 23 showed significant relationships (p-value $< 0.05$) between trophic positions and the log transformed concentrations (see Figure IV 2). Among these, 14 had TMF values $> 1$ (see Table 6.2). Nine out of those 14 analytes are alkylated polycyclic aromatic compounds, and their biomagnification can be explained based on their higher logKow values compared to HPACs and PAHs the observation can be rationalized. Trophic magnification factors were also evaluated for

individual PAC subclasses; among them, only $\sum_4$PAOHs exhibited significant biodilution, while the remaining four subclasses showed no significant trophic relationships.

However, significant TMFs should be interpreted with caution given the substantial variation in concentrations within species and the fact that the correlation between log-transformed analyte concentrations and trophic position (i.e., $R^2$ values) was consistently low. Modest $R^2$ values are common with field-derived TMF studies and due in part to natural ecological variability and species-specific differences in biotransformation rates. To fully assess the biomagnification potential of HPACs requires controlled laboratory studies, which can also provide biotransformation rates,[50] both would provide clarity to these results. Consideration of seasonal differences in HPAC concentrations and a focus on lower trophic level organisms would also contribute to a better understanding of food web dynamics of these chemicals.

**Table 6.2:** Polycyclic aromatic compounds that showed statistically significant TMF values ($p<0.05$) in the Lake Erie food web. Regression parameters for each analyte are also provided.

| Analyte Name | Slope of the TMF graphs | TMF values | P value | $R^2$ | LogKow |
|---|---|---|---|---|---|
| 1,2,3,4-Tetrahydrocarbazole | 4.16 | 64 | 0.00 | 0.210 | 4.02 |
| 5-Chloroindole | 3.21 | 25 | 0.02 | 0.080 | 3.25 |
| 2,3,4-Trimethylbenzothiophene | 0.96 | 3 | 0.02 | 0.090 | 4.63 |
| Naphthalene | 4.72 | 112 | 0.00 | 0.140 | 3.30 |
| Acenaphthene | 2.86 | 17 | 0.00 | 0.150 | 3.92 |
| Acenaphthylene | 2.10 | 8 | 0.03 | 0.080 | 3.94 |
| 1-Methylnaphthalene | 4.18 | 65 | 0.00 | 0.230 | 3.87 |
| 2-Methylnaphthalene | 4.13 | 62 | 0.00 | 0.180 | 3.86 |
| C2 Naphthalene | 4.90 | 135 | 0.00 | 0.260 | ~ 4.3 |
| C3 Naphthalene | 4.96 | 142 | 0.00 | 0.160 | ~ 4.8 |
| C1 Fluorene | 5.22 | 184 | 0.00 | 0.210 | ~ 4.1 |
| C4 Fluorene | 3.39 | 30 | 0.02 | 0.090 | ~ 6.5 |
| C2 Phenanthrene | 2.38 | 11 | 0.02 | 0.080 | ~ 5.1 |
| C1 Chrysene | 2.16 | 9 | 0.05 | 0.060 | ~ 6.0 |

Importantly, more than 75% of the analytes show no statistically significant relationship between TP and their log-transformed concentration (see Table IV 8). This corroborates the findings of our multivariant models that showed habitat-specific exposure and not trophic position is the primary driver of PAC accumulation.

Previous research has demonstrated that PAHs undergo biodilution resulting in lower concentrations in high tropic level species.[11,51] This effect is largely attributed to the AhR-mediated enzymatic pathway, which converts non-polar PAHs into more polar, excretable metabolites.[52] While HPACs are likely subject to similar biotransformation processes, the non-significant ($p < 0.05$) TMF values across most analytes suggest that habitat-specific exposure is a more prominent factor governing PAC concentrations in the Lake Erie food web than trophic transfer. The overwhelming dominance of non-priority PACs (accounting for over 95% of the total burden)

further underscores the need to expand environmental monitoring and risk assessment to include HPACs, whose distinct habitat-driven bioaccumulation may increase the risk of adverse biological effects in fish.

### 6.5.4 Future management of PACs in the Lake Erie ecosystem

A three-tier prioritization workflow was developed to identify PACs with the highest potential to negatively impact Lake Erie fish. Initially, compounds were filtered for a detection frequency of ≥ 80%. These candidates were then evaluated against three screening criteria: i) relative abundance (>10 of total PACs), ii) TMF values (TMF >1, p value < 0.05), and iii) structural features associated with biological activity, such as alkylation, halogenation, or the presence of heteroatoms. Compounds meeting two or more criteria were classified as Tier 1 (high priority), while those meeting one criterion were assigned to Tier 2 (moderate priority). Compounds consistently detected (≥ 80% detection frequency) that met none of the three criteria were categorized as Tier 3 (surveillance priority). Tier 1 compounds include indole, 4-phenylpyridine, 5-chloroindole, 1,2,3,4-Tetrahydrocarbazole, 2,3,4-Trimethylbenzothiophene and C2 Chrysene. A complete list of analytes assigned to each tier is available in Table 6.3.

**Table 6.3:** Three-tier prioritization pollutant framework of PACs for Lake Erie fish

| Analyte name | Mean relative abundance | Abundance ≥ 10% of total PACs | TMF > 1 and p value < 0.05 | Structure criteria | No of criteria met | Tier |
|---|---|---|---|---|---|---|
| Indole | 25.90 | TRUE | FALSE | TRUE | 2 | 1 (High Priority) |
| C2 Chrysene | 10.96 | TRUE | FALSE | TRUE | 2 | 1 (High Priority) |
| 4-Phenylpyridine | 10.69 | TRUE | FALSE | TRUE | 2 | 1 (High Priority) |
| 5-Chloroindole | 4.14 | FALSE | TRUE | TRUE | 2 | 1 (High Priority) |
| 1,2,3,4-Tetrahydrocarbazole | 1.26 | FALSE | TRUE | TRUE | 2 | 1 (High Priority) |
| 2,3,4-Trimethylbenzothiophene | 0.15 | FALSE | TRUE | TRUE | 2 | 1 (High Priority) |
| 5-Bromoindole | 4.56 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| C4 Naphthalene | 1.20 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| Quinoline | 0.89 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| Isoquinoline | 0.78 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| 2,3-Dimethylbenzothiophene | 0.73 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| Carbazole | 0.29 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| 3,6-Dimethylcarbazole | 0.27 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| 2-Methylquinoline | 0.17 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| Dibenzothiophene | 0.16 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| 2-Bromoquinoline | 0.16 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| 3-Methylphenanthrene | 0.14 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| 2,3-Dimethylindole | 0.14 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| 9/4-Methylphenanthrene | 0.13 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| 2-Methylphenanthrene | 0.12 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| 1-Methylphenanthrene | 0.10 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| 3,6-Dichlorocarbazole | 0.05 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| 2,4,7-Trimethyldibenzothiophene | 0.04 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| 2-Methylacridine | 0.03 | FALSE | FALSE | TRUE | 1 | 2 (Moderate Priority) |
| 2-Methyldibenzofuran | 3.25 | FALSE | FALSE | FALSE | 0 | 3 (Surveillance Priority) |
| Phenanthrene | 1.38 | FALSE | FALSE | FALSE | 0 | 3 (Surveillance Priority) |
| Fluorene | 0.49 | FALSE | FALSE | FALSE | 0 | 3 (Surveillance Priority) |

## 6.6 CONCLUSIONS

This study identified the distribution and patterns of 104 polycyclic aromatic compounds in nine different fish species from the western basin of Lake Erie. In most species, alkylated- and nitrogen-containing PACs account for greater than 60% of the ∑PACs burden, while the 16-priority PAHs accounted for only about 5% of ∑PACs. Low molecular weight PANHs such as indole and its halogenated derivatives and 4-phenylpyridine, and high molecular weight APAHs such as C3-fluorene, C4-fluorene and C4-chrysene were the most abundant analytes across all species. Multivariate ordination analysis revealed that HPACs profiles were more species- and habitat-specific than the homocyclic PACs. HPAC distributions clearly differentiated fish communities based on habitat and physiological (i.e., biotransformation, body size) traits, whereas PAH and APAH profiles appeared more uniform, suggesting weaker habitat-specific influences. Our study indicates distinct point sources of PACs into the Lake Erie ecosystem; however, more sophisticated source apportionment approaches that take into account differences in chemical persistence are needed to corroborate these observations. Collectively, the findings highlight the need for expanding current environmental monitoring programs beyond the traditional parent PAHs. The TMFs and BMFs results further emphasize the importance of controlled laboratory studies to evaluate the bioaccumulation, biotransformation, and toxicity potential of these underrepresented PAC subclasses, particularly those that are abundant yet poorly regulated. Incorporating HPACs into long-term monitoring and risk assessment frameworks will improve our understanding of PAC sources, fate, and ecological impacts across freshwater food webs.

## 6.7 ACKNOWLEDGEMENTS

This work was supported by a Natural Sciences and Engineering Research Council Discovery Grant to GTT (Grant number: RGPIN-2024-04947) and ATF (RGPIN-2020-05500) and a University of Manitoba Graduate Fellowship to NNV.

## 6.8 DECLARATION OF GENERATIVE AI AND AI ASSISTED TECHNOLOGIES IN THE MANUSCRIPT PREPARATION PROCESS

During the preparation of this work, the authors used OpenAI. (2026). ChatGPT and Google. (2026). Gemini in order to generate and proofread the R codes. After using these tools/services the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

## 6.9 REFERENCES


(1) Baumann, P. C.; Whittle, D. M. The Status of Selected Organics in the Laurentian Great Lakes: An Overview of DDT, PCBs, Dioxins, Furans, and Aromatic Hydrocarbons; **1988**; Vol. 11.

(2) Norstrom, R. J. Polychlorinated Dibenzo-p-Dioxins and Dibenzofurans in the Great Lakes. Handbook of Environmental Chemistry, Volume 5: *Water Pollution.* **2006**, pp 71–150. https://doi.org/10.1007/698_5_040.

(3) Gandhi, N.; Gewurtz, S. B.; Drouillard, K. G.; Kolic, T.; MacPherson, K.; Reiner, E. J.; Bhavsar, S. P. Dioxins in Great Lakes Fish: Past, Present and Implications for Future Monitoring. *Chemosphere* **2019**, 222, 479–488. https://doi.org/10.1016/j.chemosphere.2018.12.139.

(4) Gislason, D.; McLaughlin, R. B.; Robinson, B. W. Synchronous Changes in Length at Maturity across Four Species of Lake Erie Fish with Different Harvest Histories. *Canadian Journal of Fisheries and Aquatic Sciences* **2021**, 78 (6), 721–737. https://doi.org/10.1139/cjfas-2019-0408.

(5) Steinman, A. D.; Cardinale, B. J.; Munns, W. R.; Ogdahl, M. E.; Allan, J. D.; Angadi, T.; Bartlett, S.; Brauman, K.; Byappanahalli, M.; Doss, M.; Dupont, D.; Johns, A.; Kashian, D.; Lupi, F.; McIntyre, P.; Miller, T.; Moore, M.; Muenich, R. L.; Poudel, R.; Price, J.; Provencher, B.; Rea, A.; Read, J.; Renzetti, S.; Sohngen, B.; Washburn, E. Ecosystem Services in the Great Lakes. *Journal of Great Lakes Research.* International Association of Great Lakes Research June 1, **2017**, pp 161–168. https://doi.org/10.1016/j.jglr.2017.02.004.

(6) Heuvel, C. E.; Zhao, Y. M.; Fisk, A. T. Food Web Structure across Basins in Lake Erie, a Large Freshwater Ecosystem. Canadian *Journal of Fisheries and Aquatic Sciences* **2025**, 82. https://doi.org/10.1139/cjfas-2024-0028.

(7) Allinger, L. E.; Reavie, E. D. The Ecological History of Lake Erie as Recorded by the Phytoplankton Community. Journal of *Great Lakes Research.* September **2013**, pp 365–382. https://doi.org/10.1016/j.jglr.2013.06.014.

(8) Berthiaume, A.; Galarneau, E.; Marson, G. Polycyclic Aromatic Compounds (PACs) in the Canadian Environment: Sources and Emissions. *Environmental Pollution*. Elsevier Ltd January 15, **2021**. https://doi.org/10.1016/j.envpol.2020.116008.

(9) Achten, C.; Andersson, J. T. Overview of Polycyclic Aromatic Compounds (PAC). *Polycycl. Aromat. Compd*. **2015**, 35 (2–4), 177–186. https://doi.org/10.1080/10406638.2014.994071.

(10) Toxic substances list: schedule 1. Canadian Environmental Protection Act 1999, https://www.canada.ca/en/environment-climate-change/services/canadian-environmental-protection-act-registry/substances-list/toxic/schedule-1.html.

(11) Wan, Y.; Jin, X.; Hu, J.; Jin, F. Trophic Dilution of Polycyclic Aromatic Hydrocarbons (PAHs) in a Marine Food Web from Bohai Bay, North China. *Environ. Sci. Technol*. **2007**, 41 (9), 3109–3114. https://doi.org/10.1021/es062594x.

(12) Takeuchi, I.; Miyoshi, N.; Mizukawa, K.; Takada, H.; Ikemoto, T.; Omori, K.; Tsuchiya, K. Biomagnification Profiles of Polycyclic Aromatic Hydrocarbons, Alkylphenols and Polychlorinated Biphenyls in Tokyo Bay Elucidated by Δ13C and Δ15N Isotope Ratios as Guides to Trophic Web Structure. *Mar. Pollut. Bull.* **2009**, 58 (5), 663–671. https://doi.org/10.1016/j.marpolbul.2008.12.022.

(13) León Paumen, M.; de Voogt, P.; van Gestel, C. A. M.; Kraak, M. H. S. Comparative Chronic Toxicity of Homo- and Heterocyclic Aromatic Compounds to Benthic and Terrestrial Invertebrates: Generalizations and Exceptions. *Science of the Total Environment* **2009**, 407 (16), 4605–4609. https://doi.org/10.1016/j.scitotenv.2009.04.039.

(14) Stout, S. A.; Emsbo-Mattingly, S. D.; Douglas, G. S.; Uhler, A. D.; McCarthy, K. J. Beyond 16 Priority Pollutant PAHs: A Review of PACs Used in Environmental Forensic Chemistry. *Polycycl. Aromat. Compd.* **2015**, 35 (2–4), 285–315. https://doi.org/10.1080/10406638.2014.891144.

(15) Andersson, J. T.; Achten, C. Time to Say Goodbye to the 16 EPA PAHs? Toward an Up-to-Date Use of PACs for Environmental Purposes. *Polycycl. Aromat. Compd*. **2015**, 35 (2–4), 330–354. https://doi.org/10.1080/10406638.2014.991042.

(16) Niño-Zambrano, E. R.; Bahia, P. V. B.; Nascimento, M. M.; Sampaio, F. X. A.; Duarte, R. da S.; Hatje, V.; Machado, M. E.; de Andrade, J. B.; Rocha, G. O. da. Unconventional Polycyclic Aromatic Compounds Distribution Profiles in Sediment Cores Collected from Todos Os Santos Bay, Northeastern Brazil. *J. Hazard. Mater*. **2025**, 488. https://doi.org/10.1016/j.jhazmat.2025.137438.

(17) Manzano, C. A.; Marvin, C.; Muir, D.; Harner, T.; Martin, J.; Zhang, Y. Heterocyclic Aromatics in Petroleum Coke, Snow, Lake Sediments, and Air Samples from the Athabasca Oil Sands Region. *Environ. Sci. Technol.* **2017**, 51 (10), 5445–5453. https://doi.org/10.1021/acs.est.7b01345.

(18) Vitharana, N. N.; Halldorson, T.; Dereviankin, M.; Paddock, B.; Gerring, A.; Marvin, C.; Helm, P.; Thomas, P. J.; Idowu, I. G.; Tomy, G. T. Heterocyclic Aromatic Compounds in Lake Ontario Sediments: Spatiotemporal Distribution and New Approaches to Fingerprinting. *Science of the Total Environment* **2025**, 986. https://doi.org/10.1016/j.scitotenv.2025.179802.

(19) Metcalfe, C. D.; Metcalfe, T. L.; Riddle, G.; Douglas Haffner, G. Aromatic Hydrocarbons in Biota from the Detroit River and Western Lake Erie. *J. Great Lakes Res*. **1997**, 23 (2), 160–168. https://doi.org/10.1016/S0380-1330(97)70893-1.

(20) Wu, Y.; Tan, H.; Zhou, C.; Crimmins, B. S.; Holsen, T. M.; Chen, D. Bioaccumulation and Spatiotemporal Trends of Polyhalogenated Carbazoles in Great Lakes Fish from 2004 to 2016. *Environ. Sci. Technol.* **2018**, 52 (8), 4536–4545. https://doi.org/10.1021/acs.est.8b00427.

(21) Kimbrough, K.; Jacob, A.; Regan, S.; Davenport, E.; Edwards, M.; Leight, A. K.; Freitag, A.; Rider, M.; Johnson, W. E. Characterization of Polycyclic Aromatic Hydrocarbons in the Great Lakes Basin Using Dreissenid Mussels. *Environ. Monit. Assess.* **2021**, 193 (12). https://doi.org/10.1007/s10661-021-09401-7.

(22) Vitharana, N. N.; Halldorson, T.; Paddock, B.; Tomy, T.; Gerring, A.; Marvin, C.; Heuvel, C. E.; Fisk, A. T.; Tomy, G. T. Using Chemometrics to Discern Patterns and Spatial Distribution of Heterocyclic Aromatic Compounds in Lake Erie Fish. *Environmental Pollution* **2026**, 399, 128121. https://doi.org/10.1016/j.envpol.2026.128121

(23) Heuvel, C. E.; Zhao, Y.; Fisk, A. T. Influence of Spatial and Temporal Variation on Establishing Stable Isotope Baselines of Δ15N, Δ13C, and Δ34S in a Large Freshwater Lake. *Freshw. Biol*. **2023**, 68 (5), 806–820. https://doi.org/10.1111/fwb.14065.

(24) Idowu, I.; Francisco, O.; Thomas, P. J.; Johnson, W.; Marvin, C.; Stetefeld, J.; Tomy, G. T. Validation of a Simultaneous Method for Determining Polycyclic Aromatic Compounds and Alkylated Isomers in Biota. *Rapid Communications in Mass Spectrometry* **2018**, 32 (3), 277–287. https://doi.org/10.1002/rcm.8035.

(25) Vitharana, N. N.; Halldorson, T.; Xia, Z.; Francisco, O.; Marvin, C.; Thomas, P. J.; Liebzeit, E.; Lucas, A. M.; Moradi, V.; Tomy, G. T. A Validated Approach for Analysis of Heterocyclic Aromatic Compounds in Sediment Samples. *J. Chromatogr. A* **2024**, 1718. https://doi.org/10.1016/j.chroma.2024.464723.

(26) The Fitness for Purpose of Analytical Methods : A Laboratory Guide to Method Validation and Related Topics, Second edition, **2014**.; Magnusson, B., Örnemark, U., Eds.

(27) Vitharana, N. PCA_UMAP_for_HPACs_in_LE_foodweb.

(28) Vitharana, N. RDA_for_HPACs_in_LE_foodweb.

(29) Miranda, D. de A.; Peaslee, G. F.; Zachritz, A. M.; Lamberti, G. A. A Worldwide Evaluation of Trophic Magnification of Per- and Polyfluoroalkyl Substances in Aquatic Ecosystems. *Integrated Environmental Assessment and Management.* John Wiley and Sons Inc November 1, **2022**, pp 1500–1512. https://doi.org/10.1002/ieam.4579.

(30) Kwon, T.-D.; Fisher, S. W.; Gene, † §; Kim, W.; Hwang, H.; Kim, J.-E. Trophic transfer and biotransformation of polychlorinated biphenyls in zebra mussel, round goby, and smallmouth bass in lake erie, usa; **2006**; Vol. 25. https://academic.oup.com/etc/article/25/4/1068/7771437.

(31) Wallace, S. J.; de Solla, S. R.; Head, J. A.; Hodson, P. V.; Parrott, J. L.; Thomas, P. J.; Berthiaume, A.; Langlois, V. S. Polycyclic Aromatic Compounds (PACs) in the Canadian

Environment: Exposure and Effects on Wildlife. *Environmental Pollution*. Elsevier Ltd October 1, **2020**. https://doi.org/10.1016/j.envpol.2020.114863.

(32) Le Bihanic, F.; Morin, B.; Cousin, X.; Le Menach, K.; Budzinski, H.; Cachot, J. Developmental Toxicity of PAH Mixtures in Fish Early Life Stages. Part I: Adverse Effects in Rainbow Trout. *Environmental Science and Pollution Research* **2014**, 21 (24), 13720–13731. https://doi.org/10.1007/s11356-014-2804-0.

(33) Incardona, J. P.; Linbo, T. L.; French, B. L.; Cameron, J.; Peck, K. A.; Laetz, C. A.; Hicks, M. B.; Hutchinson, G.; Allan, S. E.; Boyd, D. T.; Ylitalo, G. M.; Scholz, N. L. Low-Level Embryonic Crude Oil Exposure Disrupts Ventricular Ballooning and Subsequent Trabeculation in Pacific Herring. *Aquatic Toxicology* **2021**, 235. https://doi.org/10.1016/j.aquatox.2021.105810.

(34) Incardona, J. P.; Carls, M. G.; Holland, L.; Linbo, T. L.; Baldwin, D. H.; Myers, M. S.; Peck, K. A.; Tagal, M.; Rice, S. D.; Scholz, N. L. Very Low Embryonic Crude Oil Exposures Cause Lasting Cardiac Defects in Salmon and Herring. *Sci. Rep*. **2015**, 5. https://doi.org/10.1038/srep13499.

(35) Pulster, E. L.; Fogelson, S.; Carr, B. E.; Mrowicki, J.; Murawski, S. A. Hepatobiliary PAHs and Prevalence of Pathological Changes in Red Snapper. *Aquatic Toxicology* **2021**, 230. https://doi.org/10.1016/j.aquatox.2020.105714.

(36) Dubiel, J.; Green, D.; Raza, Y.; Johnson, H. M.; Xia, Z.; Tomy, G. T.; Hontela, A.; Doering, J. A.; Wiseman, S. Alkylation of Benz[a]Anthracene Affects Toxicity to Early–Life Stage Zebrafish and In Vitro Aryl Hydrocarbon Receptor 2 Transactivation in a Position-Dependent *Manner. Environ. Toxicol. Chem.* **2022**, 41 (8), 1993–2002. https://doi.org/10.1002/etc.5396.

(37) Ghosh, P.; Mukherji, S. Fate, Detection Technologies and Toxicity of Heterocyclic PAHs in the Aquatic and Soil Environments. *Science of the Total Environment.* Elsevier B.V. September 20, **2023**. https://doi.org/10.1016/j.scitotenv.2023.164499.

(38) Idowu, O.; Semple, K. T.; Ramadass, K.; O'Connor, W.; Hansbro, P.; Thavamani, P. Beyond the Obvious: Environmental Health Implications of Polar Polycyclic Aromatic Hydrocarbons. *Environment International*. Elsevier Ltd February 1, **2019**, pp 543–557. https://doi.org/10.1016/j.envint.2018.12.051.

(39) C.J. van Leeuwen, T. G. V. Risk Assessment of Chemicals: An Introduction Second Edition; **2007**.

(40) Heuvel, C. E.; Haffner, G. D.; Zhao, Y.; Colborne, S. F.; Despenic, A.; Fisk, A. T. The Influence of Body Size and Season on the Feeding Ecology of Three Freshwater Fishes with Different Diets in Lake Erie. *J. Great Lakes Res*. **2019**, 45 (4), 795–804. https://doi.org/10.1016/j.jglr.2019.05.001.

(41) Idowu, O.; Tran, T. K. A.; Baker, P.; Farrel, H.; Zammit, A.; Semple, K. T.; O'Connor, W.; Thavamani, P. Bioavailability of Polycyclic Aromatic Compounds (PACs) to the Sydney Rock Oyster (Saccostrea Glomerata) from Sediment Matrices of an Economically Important Australian

Estuary. *Science of the Total Environment* **2020**, 736. https://doi.org/10.1016/j.scitotenv.2020.139574.

(42) Jake, M.; Zanden, V.; Cabana, G.; Rasmussen, J. B. Comparing Trophic Position of Freshwater Fish Calculated Using Stable Nitrogen Isotope Ratios (d 15 N) and Literature Dietary Data; **1997**.

(43) Legendre, P.; Oksanen, J.; ter Braak, C. J. F. Testing the Significance of Canonical Axes in Redundancy Analysis. *Methods Ecol. Evol*. **2011**, 2 (3), 269–277. https://doi.org/10.1111/j.2041-210X.2010.00078.x.

(44) Guo, S.; He, F.; Tang, T.; Tan, L.; Cai, Q. Intra-Annual Fluctuations Dominating Temporal Dynamics of Benthic Diatom Assemblages in a Chinese Mountainous River. *Ann. Limnol*. **2020**, 56. https://doi.org/10.1051/limn/2020020.

(45) Stow, C. A. Factois Associated with PCB Concentrations in Lake Michigan Salmo Nids; **1995**; Vol. 29. https://pubs.acs.org/sharingguidelines.

(46) Tucker, J.; Sheats, N.; Giblin, A. E.; Hopkinson, C. S.; Montoya, J. P. Using Stable Isotopes to Trace Sewage-Derived Material through Boston Harbor and Massachusetts Bay. www.elsevier.com/locate/marenvrev.

(47) Barton, D. R.; Johnson, R. A.; Campbell, L.; Petruniak, J.; Patterson, M. Effects of Round Gobies (Neogobius Melanostomus) on Dreissenid Mussels and Other Invertebrates in Eastern Lake Erie, 2002-2004. *J. Great Lakes Res.* **2005**, 31 (SUPPL. 2), 252–261. https://doi.org/10.1016/S0380-1330(05)70318-X.

(48) Muth, K. M.; Busch, W. D. N. Food of Forage Fishes in Western Lake Erie, 1975–76. *J. Great Lakes Res*. **1989**, 15 (2), 217–222. https://doi.org/10.1016/S0380-1330(89)71477-5.

(49) Knight, R. L.; Margraf, F. J.; Carline, R. F. Piscivory by Walleyes and Yellow Perch in Western Lake Erie. *Trans. Am. Fish. Soc*. **1984**, 113 (6), 677–693. https://doi.org/https://doi.org/10.1577/1548-8659(1984)113<677:PBWAYP>2.0.CO;2.

(50) Fisk, A. T.; Tomy, G. T.; Cymbalisty, C. D.; Muir, D. C. G. Dietary accumulation and quantitative structure-activity relationships for depuration and biotransformation of short (c 10 ), medium (c 14 ), and long (c 18 ) carbon-chain polychlorinated alkanes by juvenile rainbow trout (Oncorhynchus mykiss); **2000**; Vol. 19. https://academic.oup.com/etc/article/19/6/1508/7830271.

(51) Zhang, G.; Pan, Z.; Wang, X.; Mo, X.; Li, X. Distribution and Accumulation of Polycyclic Aromatic Hydrocarbons (PAHs) in the Food Web of Nansi Lake, China. *Environ. Monit. Assess*. **2015**, 187 (4), 1–12. https://doi.org/10.1007/s10661-015-4362-4.

(52) Whyte, J. J.; Jung, R. E.; Schmitt, C. J.; Tillitt, D. E. Ethoxyresorufin-O-Deethylase (EROD) Activity in Fish as a Biomarker of Chemical Exposure. *Critical Reviews in Toxicology*. Informa Healthcare **2000**, pp 347–570. https://doi.org/10.1080/10408440091159239.

# CHAPTER 7 HETEROCYCLIC AROMATIC COMPOUNDS: THE NEXT FRONTIER IN ENVIRONMENTAL FORENSIC SCIENCE

I wrote this opinion piece based on the results of my own research and that of others to argue that HPACs possess superior diagnostic capabilities compared with homocyclic PACs in environmental studies. I presented empirical evidence demonstrating that HPACs provide more effective source fingerprints for source apportionment applications and emphasized the need for the scientific community to generate persistence data for these compounds and to move beyond traditional polycyclic aromatic hydrocarbons, including HPACs, in future environmental monitoring and risk assessment studies.

This chapter has been previously published as Vitharana Nipuni N., Marvin Christopher, Thomas Philippe J., Idowu Ifeoluwa Grace, Sandau Courtney D., Dupuis-Smith Reyd., Provencher Jennifer F., Fisk, Aaron T., Tomy Gregg T, Heterocyclic Aromatic Compounds: The Next Frontier in Environmental Forensic Science, Environmental Science & Technology, 2026. Copyright © 2026. American Chemical Society. Reprinted with permission.

## 7.1 A VIEWPOINT ON HPACs

Polycyclic aromatic compounds (PACs) are an important class of organic pollutants introduced into the environment through petrogenic, pyrogenic, and biogenic processes, including natural combustion, microbial transformation, and diagenetic formation pathways. This diverse family includes polycyclic aromatic hydrocarbons (PAHs), alkylated-PAHs (APAHs), halogenated-PAHs (HPAHs), and hetero-polycyclic aromatic compounds (HPACs). Because PAC input processes, or sources, generate complex chemical mixtures rather than isolated subclasses, the presence of alkylated and heterocyclic PACs in contaminated systems is inevitable.

Over a decade ago, Andersson and Achten argued for the inclusion of alkylated, halogenated, and heterocyclic PACs into environmental assessments, warning that omitting them underestimates ecological risk by orders of magnitude.[1] Based on our work with field samples and reference materials, we have further demonstrated that HPACs are more effective diagnostic markers than PAHs for source attribution in aquatic systems, compared to their homocyclic analogues.

Several studies have investigated underrepresented PAC subclasses in the past decade using non-targeted analytical approaches.[2] These studies demonstrate that HPACs are widespread in the environment and have the potential to bioaccumulate. Despite this, HPACs have not been incorporated into environmental regulatory monitoring programs. This exclusion reflects practical challenges: a vast number of structural isomers, a scarcity of authentic standards and certified reference materials, and analytical protocols that fail to capture all PACs subclasses.[3] As a result, current risk assessments are often incomplete and may underestimate the true environmental and biological impacts of PAC mixtures.

Nitrogen, sulfur or/and oxygen containing hetero-polycyclic aromatic compounds possess at least one heteroatom in their aromatic structure imparting greater polarity than their homocyclic counterparts. Although HPACs share similar origins with PAHs, they are generally considered minor constituents and typically occur at lower concentrations in environmental matrices, particularly in abiotic samples. Recent evidence indicates that incorporating HPACs into environmental analysis provides additional insight into PAC sources, fate, and behavior.[4,5] This viewpoint advances the case that HPACs are not merely secondary analytes but essential diagnostic chemical fingerprints for resolving source inputs in spatial and species-dependent environmental studies.

The selection of suitable chemical markers is central to reliable and accurate source fingerprinting. Compared to homocyclic PACs, HPACs offer enhanced source specificity that increases their diagnostic utility.[6] Both PAHs and APAHs are produced in large quantities by a wide range of petrogenic and pyrogenic processes, including fossil fuel combustion, biomass burning, petroleum spills, and industrial activities. While general patterns exist, such as high molecular weight PAH enrichment in pyrogenic emissions and alkylated PACs in petrogenic sources, the widespread occurrence and high concentrations of homocyclic PACs often result in overlapping compositional profiles. This ubiquity complicates source attribution and can lead to ambiguous or biased interpretations.

Conversely, HPACs provide more specific, source-dependent signals. For example, azaarenes (N-PACs) are commonly associated with high-temperature combustion of nitrogen-rich fuels like coal, biomass, and diesel. Sulphur containing PACs are closely linked to S-rich petroleum sources and heavy oils whereas O-containing PACs may arise from combustion but are also formed through

oxidative weathering and microbial transformation of parent PAHs. Unlike homocyclic PACs, the unique characteristics of HPACs enable them to unambiguously differentiate combustion-derived inputs, petroleum sources, and secondary formation pathways, such as microbial transformation.

Beyond source specificity, in order for fingerprinting markers to be effective, they must also exhibit environmental persistence.[7] Evidence suggests that focusing on persistent HPACs substantially improves fingerprinting, whereas profiles based solely on the 16 priority PAHs show little change even after non-persistent compounds are removed. This difference likely reflects the high abundance and environmental ubiquity of PAHs, which can dominate chemometric analyses and obscure more subtle source-specific signals. In contrast, persistent species of HPACs, despite their lower concentrations, retain distinct and informative compositional patterns.[5] Nevertheless, broader application of HPACs in source fingerprinting remains limited by the scarcity of experimental data on their persistence and transformation, particularly in biological organisms. Addressing this knowledge gap through targeted laboratory and field studies will further strengthen the forensic usefulness of HPACs.

These advantages of HPACs for fingerprinting become even more apparent in biological matrices. In biota, HPACs often constitute a higher proportion of the total PAC burden than in abiotic matrices, likely due to their increased polarity and enhanced bioavailability. In biological samples, HPAC profiles frequently exhibit pronounced species- and location-specific patterns that are less evident for homocyclic PACs.[8] This specificity in organisms enhances their value for investigating contaminant sources by providing insights into biological exposure.

Location-specific HPAC patterns reflect both source and differences in environmental processing. Structural diversity among HPACs may result in variable photochemical degradation, microbial

metabolism, and abiotic transformation, causing persistence and composition to depend strongly on local conditions such as oxygen availability, sediment characteristics, organic carbon content, temperature, and microbial community structure.[9] These factors can produce distinct, site-specific HPAC fingerprints, convenient for assessing where organisms accumulate these contaminants but could complicate assessing specific sources. As stressed above, more controlled laboratory and field experiments to understand accumulation and biotransformation will improve our understanding of HPAC fate.

Variations in HPACs profiles across organisms further underscores their diagnostic potential. The presence of heteroatoms increases polarity and reactivity, influencing uptake mechanisms, biotransformation rates, and tissue distribution across organisms. Differences in biotransformation capacity, lipid content, and feeding strategies across biological species result in divergent HPAC absorption, distribution, metabolism, and excretion patterns.[10] In contrast, homocyclic PACs often show more uniform bioaccumulation patterns driven largely by hydrophobic partitioning. This species-specific behavior enables HPACs to serve as sensitive indicators of exposure pathways, trophic transfer, and metabolic transformation processes in aquatic and terrestrial food webs.

Although experimental validation remains limited, the evidence is clear: integrating HPACs into environmental monitoring programs would fundamentally enhance our understanding of the fate and behavior of these complex chemical mixtures. Shifting away from a PAH-only approach and incorporating HPACs into current monitoring programs eliminates much of the ambiguity in source attribution and enhances the ecological relevance of risk assessments. We call on the scientific community and regulatory bodies to move beyond legacy markers and systematically adopt HPACs as a core forensic standard for environmental assessment of PACs.

## 7.2 ACKNOWLEDGEMENTS

This work was supported by a Natural Sciences and Engineering Research Council Discovery Grant to GTT (Grant number: RGPIN-2024-04947). A.T.F. acknowledges support from the Canada Research Chairs Program (CRC-2023-00089). P.T. acknowledges support from Environment and Climate Change Canada (ECCC) and the Oil Sands Monitoring (OSM) Program (296243 PBL5 A-001103.072 2903). While this work was supported in part by the OSM program, it does not necessarily reflect the official position of the program.

## 7.3 REFERENCES


(1) Andersson, J. T.; Achten, C. Time to Say Goodbye to the 16 EPA PAHs? Toward an Up-to-Date Use of PACs for Environmental Purposes. *Polycycl. Aromat. Compd*. **2015**, 35 (2–4), 330–354. https://doi.org/10.1080/10406638.2014.991042.

(2) Machado, M. E.; Nascimento, M. M.; Bomfim Bahia, P. V.; Martinez, S. T.; Bittencourt de Andrade, J. Analytical Advances and Challenges for the Determination of Heterocyclic Aromatic Compounds (NSO-HET) in Sediment: A Review. *TrAC - Trends in Analytical Chemistry*. Elsevier B.V. May 1, **2022**. https://doi.org/10.1016/j.trac.2022.116586.

(3) Vitharana, N. N.; Halldorson, T.; Xia, Z.; Francisco, O.; Marvin, C.; Thomas, P. J.; Liebzeit, E.; Lucas, A. M.; Moradi, V.; Tomy, G. T. A Validated Approach for Analysis of Heterocyclic Aromatic Compounds in Sediment Samples. *J. Chromatogr. A* **2024**, 1718. https://doi.org/10.1016/j.chroma.2024.464723.

(4) Chibwe, L.; Manzano, C. A.; Muir, D.; Atkinson, B.; Kirk, J. L.; Marvin, C. H.; Wang, X.; Teixeira, C.; Shang, D.; Harner, T.; De Silva, A. O. Deposition and Source Identification of Nitrogen Heterocyclic Polycyclic Aromatic Compounds in Snow, Sediment, and Air Samples from the Athabasca Oil Sands Region. *Environ. Sci. Technol.* **2019**, 53 (6), 2981–2989. https://doi.org/10.1021/acs.est.8b06175.

(5) Vitharana, N. N.; Halldorson, T.; Dereviankin, M.; Paddock, B.; Gerring, A.; Marvin, C.; Helm, P.; Thomas, P. J.; Idowu, I. G.; Tomy, G. T. Heterocyclic Aromatic Compounds in Lake Ontario Sediments: Spatiotemporal Distribution and New Approaches to Fingerprinting. *Science of the Total Environment* **2025**, 986. https://doi.org/10.1016/j.scitotenv.2025.179802.

(6) Zhang, X.; Bowman, D. T.; Diamond, M. L.; Helm, P.; Jobst, K. J.; Hao, C.; Kleywegt, S.; Zhang, Z. F.; Marvin, C.; Zhang, X. Contribution of Coal Tar Sources to Polycyclic Aromatic Compounds and Associated Ecological Risk in Lake Ontario Sediments: Inference from a Novel Marker. *Environ. Sci. Technol*. **2025**. https://doi.org/10.1021/acs.est.4c10360.

(7) Zou, Y.; Wang, L.; Christensen, E. R. Problems in the Fingerprints Based Polycyclic Aromatic Hydrocarbons Source Apportionment Analysis and a Practical Solution. *Environmental Pollution* **2015**, 205, 394–402. https://doi.org/10.1016/j.envpol.2015.05.029.

(8) ’Liebzeit, E. ’Vitharana N. ’Halldorson T. ’Provencher, J. ’Mallory, M. ’Tomy, G. ’Dupuis-S. R. Species- and Tissue-Specific Accumulation of Polycyclic Aromatic Compounds in Three Arctic Seabirds. *Environ. Sci. Technol.* **2026**.

(9) Çelik, G.; Stolte, S.; Müller, S.; Schattenberg, F.; Markiewicz, M. Environmental Persistence Assessment of Heterocyclic Polyaromatic Hydrocarbons – Ultimate and Primary Biodegradability Using Adapted and Non-Adapted Microbial Communities. *J. Hazard. Mater.* **2023**, 460. https://doi.org/10.1016/j.jhazmat.2023.132370.

(10) Çelik, G.; Healy, S. A.; Stolte, S.; Mayer, P.; Markiewicz, M. Daphnia Magna as an Alternative Model for (Simultaneous) Bioaccumulation and Chronic Toxicity Assessment─Controlled Exposure Study Indicates High Hazard of Heterocyclic PAHs. *Environ. Sci. Technol.* **2025**, 59 (18), 8984–8996. https://doi.org/10.1021/acs.est.5c00384.

# CHAPTER 8 COMPREHENSIVE DATA FILTERING APPROACHES FOR NON-TARGETED ANALYSIS OF POLYCYCLIC AROMATIC COMPOUNDS

Previous chapters of this research are preliminary based on the targeted analysis of PACs. However, based on the chemical complexity of this chemical class it is impossible to develop a targeted method covering all the PACs. As argued in Chapter 7, HPACs are superior diagnostic markers for source apportionment relative to their homocyclic counterparts, yet their adoption as a forensic standard remains constrained in part by the sheer number of theoretically possible structural candidates. Therefore, ultimately non targeted analysis is required to identify the total PACs in the complex environmental samples such as Lake Ontario sediment and Lake Erie biota.

In this chapter, I present my work on non-targeted analyses. Specifically, I developed a multi-step data filtering workflow to identify PACs in highly complex environmental matrices. Each step of the data filtering process was based on first principles, which I present in careful detail. I validated the workflow using an analytical standard solution containing known PACs and then applied the workflow to identify new, previously unreported PACs in sediment and fish from Lake Ontario and Lake Erie, respectively.

A version of this chapter has been submitted as Nipuni N. Vitharana, Thor Halldorson, Brett Paddock, Thane Tomy, Chris Marvin, C.E. Heuvel, Aaron T. Fisk, Bernard Crimmins, Gregg T. Tomy; Comprehensive Data Filtering Approaches for the Non-Targeted Analysis of Homocyclic, Heterocyclic and Halogenated Polycyclic Aromatic Compounds, Environmental Science & Technology, 2026.

## 8.1 ABSTRACT

Non-targeted analysis (NTA) using comprehensive two-dimensional gas chromatography coupled with high-resolution time-of-flight mass spectrometry (GC×GC-HRTOF-MS) enables broad chemical characterization of polycyclic aromatic compounds (PACs) in complex environmental matrices, though the resulting datasets require advanced post-acquisition processing. Here, we present and validate a series of data-filtering workflows for the selective identification of polycyclic aromatic hydrocarbons (PAHs), alkylated PAHs, halogenated PAHs, and hetero-PACs. The approach integrates (i) library matching against an in-house database of 135 authentic standards, (ii) isotopic pattern-based recognition for chlorinated and brominated PACs, and (iii) an evidence-based scoring system for heterocyclic PAC detection, with results integrated into a final prioritization workflow. When the complete workflow was applied to the analytical standard, it identified all analytes, and the halogen filtering workflow identified all 28 halogenated PAC analytes. Furthermore, when applied to Lake Erie fish and Hamilton Harbour sediments, the approach reduced *ca*. 15,000 initial chromatographic peaks to *ca*. 500 and identified both targeted and non-targeted PACs. The workflow revealed previously unmonitored PAC classes, including halogenated dibenzofurans and carbazoles, not captured by our targeted analyses. Overall, our workflows substantially expand the detectable PAC chemical space while maintaining analytical rigor, providing a reproducible and transferable strategy for NTA-based environmental monitoring.

## 8.2. INTRODUCTION

Polycyclic aromatic compounds (PACs) are a complex class of organic pollutants that have two or more fused aromatic rings in their structure. While they are primarily composed of carbon and hydrogen, many PACs also contain heteroatoms such as nitrogen, sulfur, and oxygen as well as halogens. Based on their elemental composition and structural characteristics, PACs can be classified into several subclasses, including polycyclic aromatic hydrocarbons (PAHs), alkylated polycyclic aromatic hydrocarbons (APAHs), halogenated polycyclic aromatic hydrocarbons (HPAHs), and hetero-polycyclic aromatic compounds (HPACs). These compounds originate from petrogenic, pyrogenic, and biogenic processes, and occur as complex mixtures in the environment.[1,2]

To date, research has focused predominantly on the 16 United States Environmental Protection Agency priority PAHs, despite growing evidence that other classes of PACs show high abundance, persistence, bioaccumulation, and toxicity.[3,4] This oversight is largely driven by several challenges, including the high structural diversity and the large number of theoretically possible isomers, which limit the availability of commercial standards and certified reference materials.[5] Furthermore, the development of robust prioritization strategies remains constrained by a lack of toxicological data. Moreover, the scarcity of authentic analytical standards, combined with the inherent complexity of environmental matrices, continues to hinder the advancement of analytical approaches to accurately measure these compounds.[6]

Within crude oil, APAHs account for *ca*. 98% of the total PAH mass, resulting in their higher predominance in environmental samples.[7] However, the high degree of isomeric complexity presents significant analytical challenges; for example, increasing alkyl substitution on

phenanthrene from C1 to C4 increases the number of possible isomers from five to 110.[8] As a result, targeted quantitative analysis of APAHs commonly relies on the 'cluster method',[9] which can lead to overestimation of concentrations in environmental samples. In contrast, targeted analysis of HPAHs and HPACs remains limited, although recent evidence has confirmed their presence across key Canadian environments, including the Great Lakes,[10,11] Alberta Oil sands,[12,13] and the Arctic.[14]

Advances in chromatographic separation and high-resolution mass spectrometry have driven the increasing application of non-targeted analysis (NTA), particularly using comprehensive two-dimensional gas chromatography coupled with high-resolution mass spectrometry (GC×GC-HRMS).[15] Over the past decade, NTA has enabled the characterization of a broad range of PACs, including alkylated, halogenated, and heterocyclic derivatives in complex environmental matrices.[16–19] However, these approaches often generate large and complex datasets, introducing substantial uncertainty in compound identification and quantification, as well as significant data redundancy that necessitates demanding and complex data processing workflows.

Current NTA workflows typically rely on vendor-specific software and commercial spectral libraries (e.g., NIST), which have improved peak detection, deconvolution, and compound identification. Nevertheless, GC×GC-HRMS analyses frequently yield thousands of features per sample, making downstream data processing and interpretation both computationally demanding and time-intensive. Moreover, isolating PACs from a vast background of co-occurring organic contaminants remains a major challenge, underscoring a scientific gap between high-throughput data acquisition and the efficient, selective identification of compounds.[15]

In parallel, several computational and machine-learning-based approaches have emerged to accelerate non-targeted PAC-related discovery. PAH-Finder, developed by 2025, Zhang et al.[20] uses a random forest classifier trained on characteristic EI fragmentation patterns together with doubly charged fragment ions to identify candidate PAHs and derivatives in GC-HRMS data. ChloroDBPFinder applies machine learning to detect chlorinated disinfection by-products from LC-HRMS data.[21] The platform FIDDLE uses deep learning to predict molecular formula from Orbitrap and quadrupole time-of-flight instruments.[22] However, these tools were developed for specific compound classes, ionization modes, and analytical platforms. None were designed for comprehensive, non-targeted screening of homocyclic, heterocyclic, and halogenated PACs in complex environmental samples analyzed by EI GC×GC-HRMS.

To address these challenges, this study develops and validates a series of post-acquisition data-filtering workflows for selective identification of PAHs, APAHs, HPAHs, and HPACs from GCxGC-HRMS datasets. By integrating library matching, isotopic pattern recognition, and structure-based filtering, the framework improves the confidence, selectivity, and efficiency of PAC identification in complex environmental matrices. Application to Great Lakes sediment and biota samples demonstrates the ability of this approach to systematically characterize diverse PAC classes from highly complex datasets, bridging a critical gap between high-throughput NTA and reliable compound identification.

## 8.3. MATERIALS AND METHODS

### 8.3.1 Chemicals

All PAHs, APAHs, and isotopically labeled PAHs were obtained from Chiron (AS, Stiklestadveien 1, 7041 Trondheim, Norway). HPAC standards were purchased from both Chiron and Sigma-

Aldrich Canada Ltd. (Oakville, Ontario, Canada). High-purity (Optima grade) organic solvents, along with silica gel (923 grade, 100–200 mesh), alumina (60–325 mesh), anhydrous sodium sulfate, and diatomaceous earth (DE), were sourced from Fisher Chemicals (Ottawa, Canada). Details of the targeted analytes and internal standards, including their codes and CAS numbers, are provided in Table V.1. The recovery internal standards (RIS) used were naphthalene-$d_8$, fluorene-$d_{10}$, phenanthrene-$d_{10}$, pyrene-$d_{10}$, benz[a]anthracene-$d_{12}$, chrysene-$d_{12}$, benzo[b]fluoranthene-$d_{12}$, benzo[a]pyrene-$d_{12}$, and benzo[g,h,i]perylene-$d_{14}$. Anthracene-$d_{10}$ was used as the instrument performance internal standard (IPIS).

### 8.3.2 Environmental samples

Two sediment samples from two different sampling sites in Hamilton Harbour of Lake Ontario were collected in 2014. More information about sample collection and sampling sites can be found in Idowu et al.[23] Two different fish species were collected from the Western basin of Lake Erie from May to October 2019. The protocols followed during sample collection and storage are described in Heuvel et al. [21] and the Appendix V. Table V 2 depicts the sample names, sampling year, and their geographical locations.

### 8.3.3 Sample purification and extraction

Sediment and fish samples were extracted using accelerated solvent extraction (ASE) using a 1:1 mixture of DCM and ethyl acetate as the extraction solvent. Internal standards were spiked into the samples prior to the extractions. The final sample volumes were 1 mL for sediments and 200 µL for biota. The full procedures of the sample extraction and purification can be found in our previous publications and Appendix V.[6,24] The extraction and purification methods were previously validated according to the Eurachem Guide -The Fitness for the Purpose of Analytical Methods.[25]

Diatomaceous earth was processed as the field blanks for the sediment samples. Three method blanks were processed with each sample set. RIS recoveries were calculated for all samples (mean ± SD) using anthracene-$d_{10}$ as the IPIS.

### 8.3.4 Gas chromatography mass tandem spectrometry analysis

Quantitative targeted analysis was carried out at nominal mass resolution using an Agilent 7890 GC coupled with a triple quadrupole Agilent 7000C mass spectrometer (GC-MS/MS) with an electron ionization (EI) source. The instrument parameters and details of the multiple reaction monitoring methods are described by Idowu et al.[9] and Vitharana et al.[6]

### 8.3.5 GC x GC high resolution time of flight mass spectrometry analysis

High mass resolution was used for analyte confirmation and performed using an Agilent 7890A GC coupled with a Pegasus GC-HRT 4D (LECO, St Joseph, MI, USA) high-resolution time-of-flight mass spectrometer (TOF-MS) using a previously developed and validated method.[23] Injections of 2 µL were carried out at 300 °C in splitless mode. A Rxi -5Sil MS (60 m × 0.25 mm, 0.25 µm) (Restek, Bellefonte, PA, USA) column was used as the primary column, and a Rxi-17Sil MS (2 m × 0.25 mm × 0.25 µm) (Restek, Bellefonte, PA, USA) column was used as the secondary column with an SGE micro union connection. Helium was used as the carrier gas. The oven was programmed from 55 °C (held for 2 min), heated to 110 °C at 10 °C/min, further to 210 °C at 3 °C/min, and finally to 300 °C at 8 °C/min (held for 13 min). The secondary oven and modulator were kept at 5 °C and 15 °C above the primary oven temperature, respectively. The mass spectrometer was run at a resolving power (full-width half height) of 25,000 based on the peak width of *m/z* 218.9851 of perfluorotributylamine (PFTBA) in EI mode. Average mass accuracy for five consecutive tunes was –0.25 ± 0.08 ppm for 218.99 m/z value. During the chromatographic

run, PFTBA was continuously bled into the ion source to serve as an internal standard for instrument mass calibration. Solvent blanks were run between every five injections.

### 8.3.6 Data analysis

High-resolution TOF data analysis was first conducted using the Leco ChromaTOF software (version: 5.10.59). All the parameters used in the ChromaToF data processing method can be found in the Appendix V. Chromatographic and mass spectral data were then exported as .csv files and processed using the custom-developed workflows using RStudio (version: 2026.04.0 "Globemaster Allium" released on 2026.04.18) for Windows. The R codes were optimized with the assistance of artificial intelligence (AI) tools: Anthropic. (2025; Claude 3.5 Sonnet), OpenAI (2026; ChatGPT) and Google (2026; Gemini). All raw data and R scripts can be found on GitHub (https://github.com/NipuniV/NTA-PAC-Workflow.git). The exported files from ChromaTOF were first processed through four independent filtering workflows separately, and the results were subsequently compared to identify peaks with the highest probability of being PACs. The first workflow (WF1) incorporated the GCxGC-HRTOF-MS data for the 135 authentic standards (Table V 1), which served as a PAC reference library and enabled the positive identification of targeted PACs and their isomers in the samples. Workflow 2 (WF2) used halogen isotopic patterns and exact masses of potential halogenated PACs to identify the peaks corresponding to halo-PACs. Workflow 3 (WF3) incorporated a series of filtering steps, including mass defects and exact masses of potential heteroatom-containing PACs, to identify potential HPAC peaks. Chromatographic and mass spectral data for each sample were analyzed with all three workflows, and the resulting datasets were integrated and compared in a fourth and final workflow (WF4) to assign PAC identities with higher confidence.

Further data processing steps, including blank correction, replicate analysis, and semi-quantification, were performed in Microsoft Excel. Analytes with peak areas less than 50% of the corresponding average blank peak areas were excluded. Additionally, analytes not detected in at least two out of three replicates were removed to minimize false positives. Finally, sediment samples were semi-quantified using concentration data obtained from GC–MS/MS analysis of targeted compounds. Relative abundances were estimated based on the concentrations of the closest eluting targeted compounds, and independent of MS structural features. Figure 8.1 depicts the complete methodology used for NTA of PACs in this study.

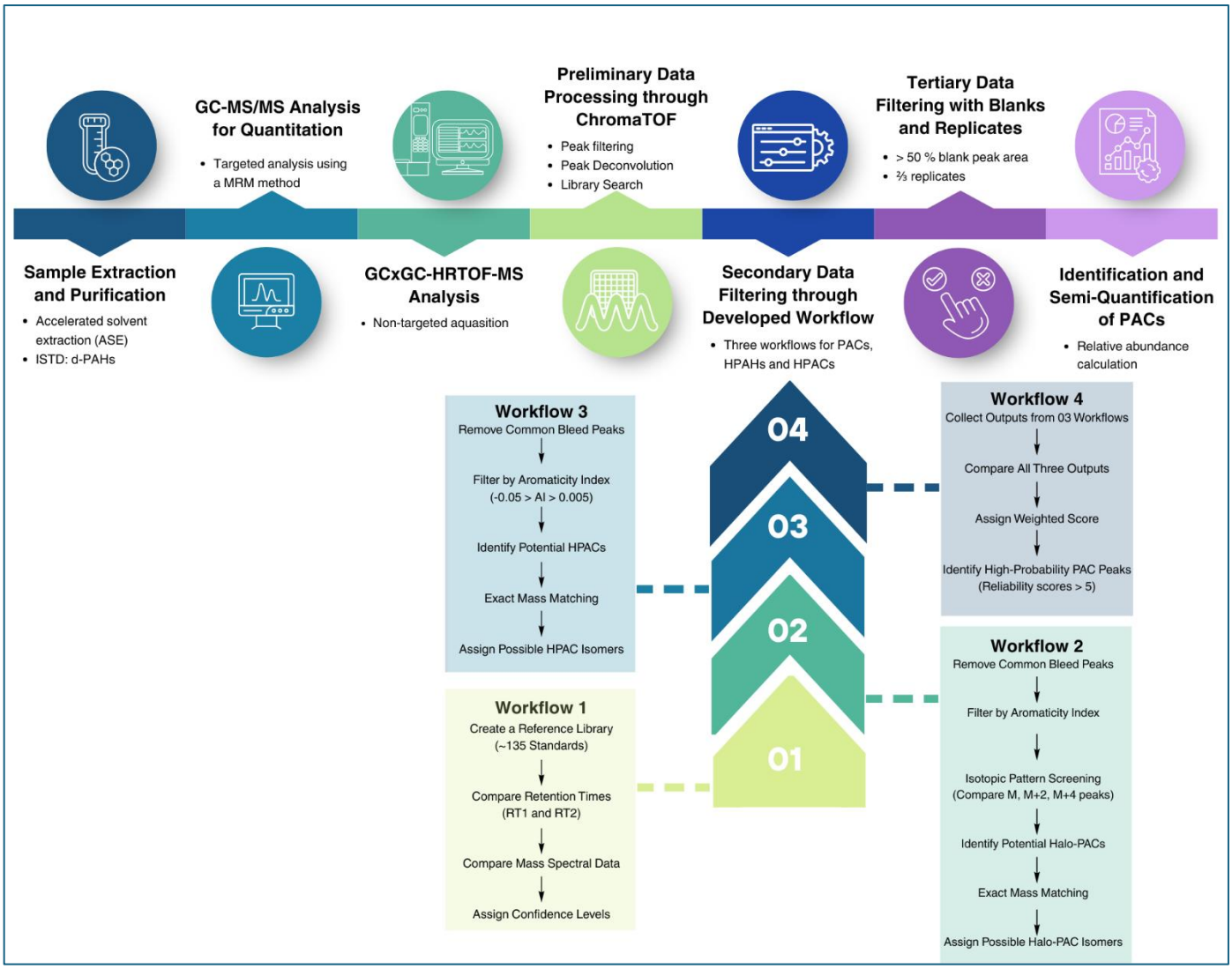



**Figure 8.1:** Complete non-targeted data analysis methodology used to identify polycyclic aromatic compounds

## 8.4. DISCUSSION

Following the identification of ~120 targeted compounds in Lake Erie fish and Lake Ontario sediments using GC-MS/MS analysis,[10,11,24] we further explored non-targeted PACs. However, in addition to matrix complexity, PACs are often present with other organic contaminants such as polychlorinated biphenyls (PCBs) that are coextracted with PACs due to their similar chemical properties. To address these challenges, multiple filtering strategies were integrated into a unified framework, improving the selectivity and confidence of PAC identification. The following sections provide a detailed description of the preliminary and secondary filtering steps and workflows.

### 8.4.1. Preliminary data filtering through ChromaTOF

A signal-to-noise (S/N) ratio threshold of 10 was applied to all chromatographic peaks in our standards, samples, and blanks. This threshold was selected to accommodate the previously observed low concentrations of targeted compounds and the reduced detectability of the HRTOF system compared to GC-MS/MS. Library matches were performed against both the NIST database and an in-house library using a stringent mass accuracy threshold of ±3 ppm, with compounds accepted with a minimum similarity score of 750. Elemental formula assignments were performed at a mass accuracy of ±5 ppm, while constraining elemental compositions and double bond equivalence (DBE) values to ranges consistent with PACs (see below).

This approach resulted in *ca*. 5000 - 15000 chromatographic peaks in sediment and biota samples, the majority of which remained unidentified after library matching. For comparison, preliminary data processing of our analytical standard solution using the chosen S/N threshold resulted in ca. 350 peaks, of which 50% were unassigned. It should be stated that the S/N of our analytes was typically two orders of magnitude greater than the extraneous peaks identified by the initial filtered

step described above. The unidentified peaks were further evaluated through manual inspection by comparing first-dimension retention times ($rt_1$) and mass spectral data with those obtained from GC–MS/MS analysis. Following this step, all 118 native standards and 17 internal standards were positively identified and isolated from the peak tables. The curated data were then exported as .csv files, with peak names, assigned elemental formulas, retention times in both dimensions, base peak masses, peak areas and normalized mass spectral data.

### 8.4.2. Secondary data filtering through customized workflows

Four independent filters were developed and incorporated into a R project (Link: https://github.com/NipuniV/NTA-PAC-Workflow.git) for the selective identification of PAC. Each workflow employed unique filtering procedures, which will be discussed in greater detail in the following sections. Finally, we use these filtering workflows to report on PACs in samples from the Great Lakes.

### 8.4.3. Workflow 1: assignment of confidence levels based on a custom-created library data of PACs

The first step of this workflow involved creating a reference in-house library file. As described above, a standard file was exported following ChromaTOF library matching and manual peak assignments. This file was reorganized and incorporated into the workflow, including the chromatographic and mass spectral characteristics of each native standard and mass-labeled internal standard (Table V 3). A mass accuracy threshold of ±5 ppm was applied, and two retention time windows were defined as ± 2 seconds for the first dimension (rt1) and ± 0.2 seconds for the second dimension (rt2). These retention time tolerances were determined by taking the average retention times plus 1 standard deviation of three standard injected on separate days and

interspersed with sample injections. Subsequently, the retention times and the four most intense mass spectral peaks were compared against the corresponding data in the library.

Four confidence levels were then assigned as follows. Compounds matching both retention times and at least three of the four most intense mass spectral peaks were assigned to Level 1 (confirmed compound). Level 2 (probable compounds) included those with matching retention times and mass spectral agreement for one to three *m/z* values. Level 3 was assigned to compounds exhibiting three or more matching ions within the specified mass accuracy threshold but with retention times that did not align with the library data. This level identifies isomers and hence the non-target compounds. Compounds with fewer than 3 matching *m/z* values and with different retention times from the library were classified as Level 4 (candidates). To avoid redundancy, peaks assigned to Level 1 or 2 were not reclassified into lower confidence levels.

According to a 2002 European Union decision report, compound identification can be supported by as few as two high-resolution mass fragments.[27] However, in the present workflow, confidence levels were assigned using a more stringent, precautionary approach to minimize false positives. Four repeated GC-injections of analytical standards were processed through WF1. With the exception of 6-methylbenzo[a]pyrene and 1,8-dimethylphenanthrene, all native standards and internal standards were identified in at least one injection and more than 83% of analytes were consistently identified across all four injections. For comparison, only 42 could be identified using the LECO ChromaToF library search, with seven of these misassigned to incorrect isomers. The resulting output from a standard solution processed through WF1 is summarized in Table V 4.

All environmental samples were processed using this workflow, resulting in a substantial reduction in the number of detected peaks. For example, a yellow perch sample showed *ca*. 15,000

chromatographic peaks in the initial ChromaTOF output; this was reduced to 88 peaks following WF1 processing; 32 peaks were classified as Level 1 and 2, while 25 were identified as isomers in the library data (Level 3).

### 8.4.4. Workflow 2: identification of potential halogenated PACs

The second workflow was developed specifically to identify chlorinated and brominated aromatic compounds. To differentiate halo-PACs from other halogenated species and instrumental artifacts, such as GC-column bleed, a multi-stage filtering protocol was implemented. First, low-molecular weight fragments (base peak < *m/z* 100) and early eluting components (rt1<850 seconds) were excluded, as these all fell below the threshold for the target halo-PACs. Subsequently, peaks corresponding to known siloxane-based GC column bleed ions (m/z 73.0474, 207.0325, 221.0482, 253.0200, 267.0536, 281.0693, 355.0851, 429.1008) were filtered using a mass accuracy tolerance of ±10 ppm.

Next, an expanded halogen mass filter, adapted from Fernando *et al.*[17] was integrated into the workflow to identify Cl- and Br-containing candidates based on their characteristic $^{35}Cl/^{37}Cl$ and $^{79}Br/^{81}Br$ isotopic signatures. The halogen filter evaluates isotopic patterns in mass spectra of the monoisotopic peak (M) and the corresponding M+2 and M+4 mass offsets arising from halogen isotopes.

Halogenated candidates were initially screened by detecting the presence of an M+2 peak. The algorithm then searches for a corresponding monoisotopic peak (M) at a mass difference of *ca*. 1.997 Da, representative of the isotope shifts of both Cl and Br. Although element-specific exact mass differences were tested for Cl (1.9970) and Br (1.9979) independently, this approach proved overly restrictive, resulting in the exclusion of valid candidates.

Workflow 2 then evaluated the intensity ratio between the monoisotopic peak (M) and the M+2 peak, applying a minimum threshold (M/M+2 ≥ 0.40) to retain only chemically plausible isotope pairs. This criterion helps eliminate false positives that arise from unrelated ion fragments that do not reflect true halogen isotope patterns.

To differentiate between mono- and di-halogenated compounds, the workflow additionally searches for an M+4 peak corresponding to two halogen substitutions. The presence and relative intensity of this peak provide key diagnostic information. For example, chlorine exhibits a ~3:1 (M:M+2) ratio for mono-chlorinated compounds and a ~9:6:1 (M:M+2:M+4) pattern for di-chlorinated species, whereas bromine shows ~1:1 and ~1:2:1 patterns, respectively. These known isotope signatures were used to calculate M/M+2 and M/M+4 ratios, which form the basis of halogen classification. Based on these ratios, each detected isotope pattern was categorized into classes such as mono-chlorinated, mono-brominated, di-chlorinated, di-brominated, or mixed/poly-halogenated compounds. This classification was implemented using defined ratio thresholds, allowing systematic and repeatable annotation of halogen types. Those thresholds are summarized in Table 8.1.

**Table 8.1:** Halogen class assignment criteria used in workflow 2

| Halogen Class | M / M+2 ratio | M+4 availability | M / M+4 ratio | Chemical basis |
|---|---|---|---|---|
| Mono-Cl | ≥ 2.0 | No | > 20 | $^{35}Cl$: $^{37}Cl$ = 3: 1, M+4 is absent |
| Mono-Br | 0.7 – 1.5 | No | > 20 | $^{79}Br$: $^{81}Br$ = 1: 1, M+4 is absent |
| Di-Cl | 1.0 – 2.0 | Yes | 5 - 20 | $^{35}Cl$: $^{35}Cl^{37}Cl$: $^{37}Cl$ = 9: 6: 1 |
| Di-Br | 0.3 – 0.7 | Yes | 0.5 – 4.0 | $^{79}Br$: $^{79}Br^{81}Br$: $^{81}Br$ = 1: 2: 1 |
| Poly/mixed | Any other | Yes | Any other | Doesn't fit the above patterns |

As explained in Fernando *et al.*,[17] it was critical to remove false positives arising from natural carbon isotope patterns. Specifically, peaks that could correspond to $^{13}C$ isotopes (M+1 peaks) were excluded if their intensity exceeded 20% of the monoisotopic peak. This prevents misassignment of carbon isotopic peaks as halogen-derived M+2 signals, which is particularly important in high-resolution datasets where isotopic overlaps can occur.

Despite these extensive filtering steps, a substantial number of chromatographic peaks remained in the chromatogram for biota and sediment samples (e.g., ~1000 peaks in yellow perch). Most of these peaks were attributed to halogenated aliphatic compounds. To address this, we created an aromaticity-based filter. To selectively retain aromatic compounds, an aromaticity index was calculated from mass defect values using the steps highlighted in Figure 8.2. Instead of using the base mass or molecular mass, we applied the aromaticity index filter to M.

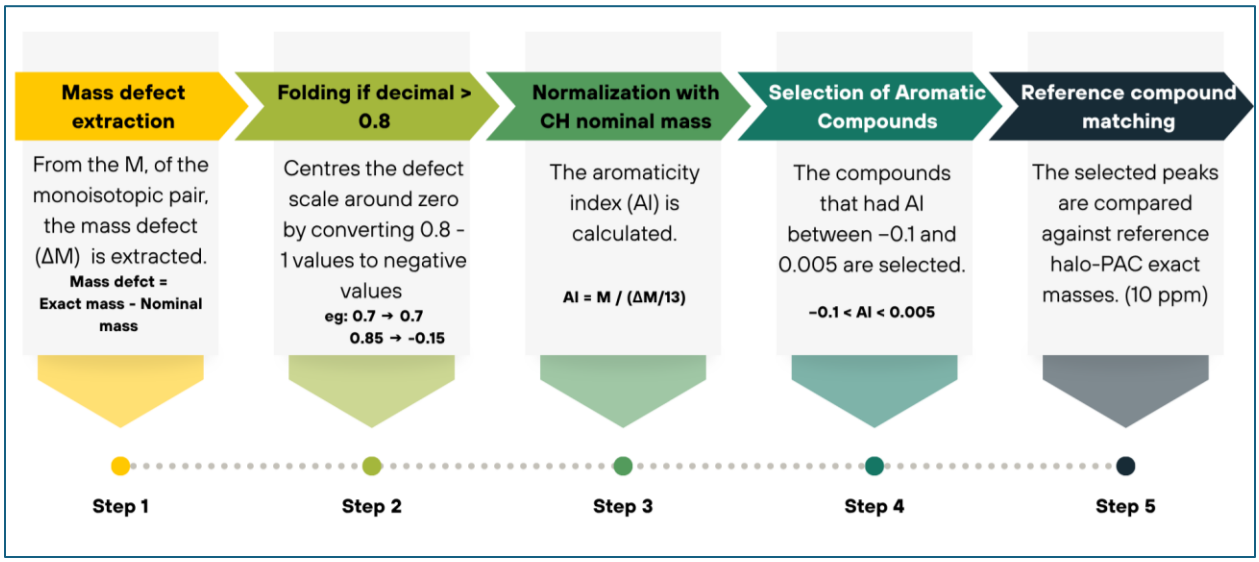



**Figure 8.2:** The steps used in the aromaticity index (AI) filter in workflow 2

The decimal portion of the monoisotopic mass was first isolated and transformed to a symmetric scale around zero to ensure consistent treatment of both positive and negative mass defects. The

adjusted mass defect was then normalized to the nominal mass of a CH group (mass defect/13), yielding the aromaticity index. For example, consider chloroquinoline ($C_9H_6ClN$), a chlorinated PANH with a monoisotopic mass of 163.01888 Da. The calculation proceeds in three steps. First, the mass defect is extracted as the decimal portion of the mass (163.01888 − 163 = 0.01888). Second, this defect is folded onto a symmetric scale. Since 0.01888 falls well below 0.8 threshold, no folding is applied. Third, the folded defect is normalized by dividing it by the mass divided by 13 by calculation aromaticity index as 0.0015. This value falls inside the acceptance window used by WF2 ($-0.05 < AI < 0.005$), and lies within the narrow band and selected as a halo-PAC.

The rationale behind this calculation is based on how different elements contribute to the mass defect. Hydrogen has a relatively large positive mass defect, while elements such as oxygen, sulfur, and especially halogens shift the mass defect in a negative direction. Nitrogen contributes positively to the mass defect, but to a lesser extent than hydrogen. Because aromatic compounds are carbon-rich and comparatively hydrogen-deficient relative to aliphatic compounds and siloxanes, they occupy a characteristic mass defect range.[28]

To define an appropriate aromaticity index range for halo-PACs, aromaticity indices were calculated for all available PAC standards (Table V 5). The highest observed value corresponded to 9-chloroflouranthene (0.0026). In the absence of halogenated naphthalene standards, a conservative cutoff of 0.005 was selected as the upper limit for halo-PACs. Applying this filter reduced the number of chromatographic peaks in the yellow perch sample from *ca*. 1000 to 500.

The acceptance window applied within WF2 is two-sided ($-0.050 < AI < 0.005$); the lower bound excludes ions with extreme negative mass defects, characteristic of heavily halogenated fragments that are not aromatic PACs, while the upper bound retains the limit described in the preceding

paragraph. Unsurprisingly, the aromaticity index filter resulted in residual peaks from other environmental pollutants such as PCBs and halogenated dioxins.

Because the aromaticity index is computed from the monoisotopic mass of candidate ions, it is a deterministic function of elemental composition. Its performance can therefore be evaluated in silico against any set of molecular formula, independently of acquired spectra. Because the aromaticity index filter in WF2 was applied only after isotope-pattern matching had restricted the candidate pool to halogen-bearing peaks, its role at that stage is specifically to retain halogenated aromatic compounds while rejecting halogenated aliphatic species.

An in-silico evaluation of the aromaticity index across a broad reference set of molecular formula was performed as a benchmark step. To do so, a reference set of 505 compounds was compiled from PubChem and NIST elemental compositions and divided into four categories: homocyclic PACs (n = 84), heterocyclic PACs bearing N, S, or O ring atoms (n = 95), halogenated compounds spanning both aromatic and aliphatic structures (n = 96), and a diverse non-PAC background set including aliphatic hydrocarbons, fatty acids and esters, siloxanes, phthalates, and phenols (n = 244). For every compound the monoisotopic exact mass was computed directly from the molecular formula, and the aromatic index was calculated using the procedure described in WF2 (Table V 6).

The 84 homocyclic PACs had a mean aromaticity index of 0.0062 (SD 0.0011; range 0.0036–0.0088), and the 95 heterocyclic PACs scored a mean of 0.0039 (SD 0.0015; range 0.0004–0.0064). Both aromatic classes were confined to a narrow positive band (0.0004–0.0088) regardless of ring number, alkylation, or heteroatom content, confirming that aromatic classes occupy a characteristic mass-defect region as described above.

The aromaticity index filter was then evaluated against our 96 halogen-bearing candidates. The acceptance window described earlier ($-0.050 < AI < 0.005$) retained 58 of 70 halogenated aromatics (sensitivity 0.83; precision 0.70). The performance characteristic calculations for all the benchmarking tests of this study are reported in Appendix V. Importantly, all halo-PAC-like compounds in the set were retained: all six chloronaphthalenes, both chlorostyrenes, and all ten PCBs fell inside the window. The twelve halogenated aromatics excluded were polybrominated diphenyl ethers and polybrominated benzenes (PBDEs, hexabromobenzene, tribromobenzene, tribromophenol), which are not PACs. Their exclusion reflects the same physicochemical logic by which polybrominated species carry a larger mass defect that lifts them above the aromatic band.

The window admitted 25 of 26 halogenated aliphatics, consisting of small chlorinated and brominated alkanes, alkenes, and organochlorine pesticides (e.g., DDT, aldrin), whose aromatic index values fall in a similar range to halogenated aromatics. The aromaticity index therefore does not, by itself, distinguish halogenated aliphatics from halogenated aromatics which is why WF2 does not rely on it alone.

To resolve this ambiguity, an additional filtering step was applied after the aromatic index filter. The exact masses of common halo-PACs (Table V 7) were compared against the most intense ion among M, M+2, and M+4. This further reduced the yellow perch sample dataset to 58 peaks. However, five confirmed PCB peaks were misclassified as trichlorinated acenaphthenes because the analysis relied on full isotopic pattern matching to the reference exact masses.

The finalized workflow successfully identified all 28 halogenated standards present in the PAC standard mixture. All the halogenated standards were assigned as halogenated PAC based on comparison with the reference masses. Table V 8 summarizes the results of the standard file

processed through WF2. However, there were 3 peaks in our standard solution that were misidentified as halo-PACs (phenanthrene-tri-Cl, 11H-benzo[a]carbazole-mono-Cl-mono-Br, and acenaphthylene-mono-Cl-di-Br), demonstrating that the WF2 required further filtering steps or manual peak inspection to avoid false positives. For comparison, of the 28 halo-PACs, only 11 were successfully identified during the initial data processing using ChromaTOF.

WF2 performance can be evaluated on two levels: halogen isotope-pattern detection, and specific compound identification. At the detection level, using the 135-compound in-house standard library (28 halogenated, 107 non-halogenated), WF2 correctly flagged all halogenated standards via automated isotope-pattern and mass comparison (sensitivity = 100%). None of the 107 non-halogenated standards were falsely flagged as halogenated, giving 100% specificity for halogen-versus-non-halogen discrimination.

The 3 misidentified peaks suggest that WF2 reliably identifies halogen isotope patterns, but it cannot independently confirm that the matched mass corresponds to the correct compound rather than a halogenated fragment of a lower-degree homologue already present in the mixture. In the absence of retention time constraints or fragmentation-based structural confirmation, mass coincidence within 10 ppm is insufficient for distinction. This limitation is inherent to any pattern-matching workflow that must balance sensitivity against specificity. Consequently, WF2 outputs should be treated as candidate lists that require subsequent manual inspection or verification using Kendrick mass defect (KMD) plots, as evidenced by the identification of false positives in the KMD analysis.

An external empirical test was then used to assess the performance of WF2 prior to the KMD analysis. High resolution EI mass spectra taken from RECETOX Exposome HR-[EI+]-MS library

(386 individual spectra)[29] was processed using WF2 to assess its performance. This library was constructed by the analysis of authentic compounds via GC-Orbitrap MS. Polycyclic aromatic character was assigned from the library's InChI and SMILES metadata using RDKit, requiring genuine ring fusion (aromatic rings sharing at least two atoms) consistent with the composition of the halo-PAC reference table used within WF2. On this basis, the MS library contains 191 halogenated compounds, of which 11 are fused halogenated PACs.

Twelve of the 191 halogenated compounds in the RECETOX library contain fluorine as their sole halogen; because $^{19}F$ is monoisotopic and has no M+2 isotopic patterns, these compounds fall outside the detectable scope of any isotope-pattern criterion.

WF2 flagged 209 of the 386 spectra as halogen-bearing, corresponding to a sensitivity of 0.932, specificity of 0.841, precision of 0.852, and F1 score of 0.890 against the halogenated/non-halogenated ground truth (Appendix V). Excluding fluorinated PACs, sensitivity rises to 0.955 and F1 to 0.898. Out of the 11 fused halogenated PACs, 10 were recovered by the halogen filter. As explained above, 1-fluoronaphthalene was not identified as a halo-PAC.

There were 31 non-halogenated compounds that were incorrectly flagged were either sulfur-containing pesticides or sulfur-containing heterocyclic aromatics. This interference is attributable to $^{34}S$, which occurs at the same nominal M+2 offset as $^{37}Cl$, and has a natural abundance of approximately 4.4%. Our intensity-ratio criteria alone cannot fully distinguish these S-compounds from mono-chlorination species. Importantly, this implies that sulfur-rich compounds are the principal interferent class for isotope-pattern halogen screening. Furthermore, this reinforces the requirement that WF2 outputs be treated as candidate lists subject to subsequent exact-mass and homologous-series verification, rather than as definitive identifications.

Following exact-mass matching against the halo-PAC reference table, 7 of the 11 fused halogenated PACs were retained (sensitivity 0.636, specificity 0.763, precision 0.073). This low precision reflects the base rate of halo-PACs in the library (2.8%) rather than a deficiency in the underlying detection logic, which achieves an F1 of 0.898 at the halogen-detection stage. In an environmental matrix where halo-PAC constitutes a larger fraction of the halogen-bearing population, the corresponding precision would be substantially greater.

Kendrick mass defect plots were generated for 28 halogenated PAC standards (see Figure V 1) to identify homologous series by $CH_2$ (nominal 14 Da, exact 14.01565 Da), representing one carbon-chain extension; Cl (nominal 34 Da, exact 33.96103 Da), representing a single chlorine-for-hydrogen substitution (H/Cl); and $Br_2$ (nominal 158 Da, exact 157.83604 Da), representing two bromine atoms. These plots were constructed from the WF2 output to visualize the organization of chemical space and to validate compound identifications against the reference library.[30] The H/Cl and $Br_2$ KMD plots mainly reflect differences in aromatic core structure, which correlate with the aromatic ring condensation along the y-axis. The KMD becomes more negative as the number of rings increases from two to five. Three distinct horizontal bands are visible in both plots, reflecting this ring-number stratification.

According to the H/Cl mass defect plots, the low molecular weight two-ring PACs such as 2,3-$Cl_2$-BT, 5-Cl-Ind, 2-Cl-Qui, and 2-Br-BT lie in the upper range (~ -0.17 to -0.21) of the KMD axes. The middle band (~-0.24 to -0.27) contains three ring PACs, such as 3,6-$Br_2$-Car, 2-Br-DBT, 5,6-$Br_2$-1,2-$H_2$-Acy, and 9-Cl-Acr; this represents a tighter cluster with high Cl and Br overlap. The lower band starts from ~ -0.3 and contains high molecular weight four-ring PACs such as 1-Br-Pyr and 4-Br-BaA.

The $CH_2$ KMD plot provides complementary structural information by organizing compounds according to differences in carbon backbone length, resulting in a diagonal trend rather than horizontal banding. With the exception of 1,6-$Br_2$-Pyr, all halogenated standards displayed positive KMD values, with Br-substituted compounds consistently plotting above Cl-substituted compounds at equivalent Kendrick Mass, a direct consequence of the larger positive mass defect contribution of bromine relative to chlorine. The negative $CH_2$ KMD observed for 1,6-$Br_2$-Pyr (KMD = −0.497) does not represent a distinguishing structural property of the compound. Kendrick mass defect is calculated as KMD = round(KM) − KM and has a discontinuity at every half-integer value of KM. A KM value just below 0.5 rounds down to the lower integer, producing a KMD approaching −0.5, whereas a value just above 0.5 rounds up to the upper integer, producing a KMD approaching +0.5. For 1,6-$Br_2$-Pyr, the $CH_2$-series Kendrick mass is 357.497, only 0.003 units below this boundary. This is nearly sevenfold closer to the discontinuity than any other standard in the training set, so the negative KMD results from the mathematical discontinuity rather than a genuine mass-defect characteristic.

This sensitivity is specific to the $CH_2$ series and does not extend to the H/Cl or $Br_2$ series because each series uses a different repeat unit. Consequently, the same molecular mass occupies a different position within the 0 - 1 Kendrick mass cycle for each series. The Kendrick mass of 1,6-$Br_2$-Pyr is only 0.003 units from the discontinuity in the $CH_2$ series, but 0.193 and 0.232 units from the corresponding discontinuities in the H/Cl and $Br_2$ series, respectively. Therefore, the compound is not flagged as anomalous in those two series. This behaviour is a recognized mathematical property of Kendrick mass defect analysis and can occur for any compound whose mass lies close to a half-integer Kendrick mass in a given series.

Across all three series, compounds sharing the same aromatic core but differing in halogen type occupied similar KMD values while shifting horizontally in Kendrick Mass, confirming that the aromatic framework governs the mass defect rather than the halogen substituent. This finding is consistent with the aromaticity index filter applied in WF2.

To address whether these KMD trends could provide quantitative evidence for unknown compound identification, linear regression models were fitted separately for each halogen class (mono-Cl, di-Cl, mono-Br, di-Br) within each of the three KMD series, yielding 12 models in total (Figure 8.3). For each model, 95% prediction intervals (PIs) were calculated. Prediction intervals were used instead of confidence intervals because they account for the variability of a new individual measurement rather than the uncertainty of the mean. For instance, where a large number of analytical standards were available for a specific compound class (e.g., di-bromo-PACs), the PI bands for that group were smaller than those where a smaller number of standards were available (e.g., di-chloro-PACs, see Figure 8.3). The results from steps 4 and 5 of the WF2 (*see* Figure 8.2) were used to construct KMD plots after blank correction. These KMD plots are shown in Figure V 1.

Compounds detected in an environmental sample were evaluated against all twelve models simultaneously. If an observed KMD fell within the 95% PI of a given class in all three KMD series, it received a "High" confidence KMD assignment; agreement in two of three series was scored as "Moderate"; one of three as "Low"; and no agreement as "Not supported by KMD". The above data obtained for a standard are summarized in Table V 9.

This multidimensional approach reduces the probability of a false positive class assignment compared to relying on any single series, because a compound would need to fall within the PI

band by chance across multiple independent chemical dimensions simultaneously. Peaks that had Kendrick Mass outside the KM range spanned by the 28 training standards were additionally flagged as extrapolations, where the widening PI and reduced training support make predictions less reliable.

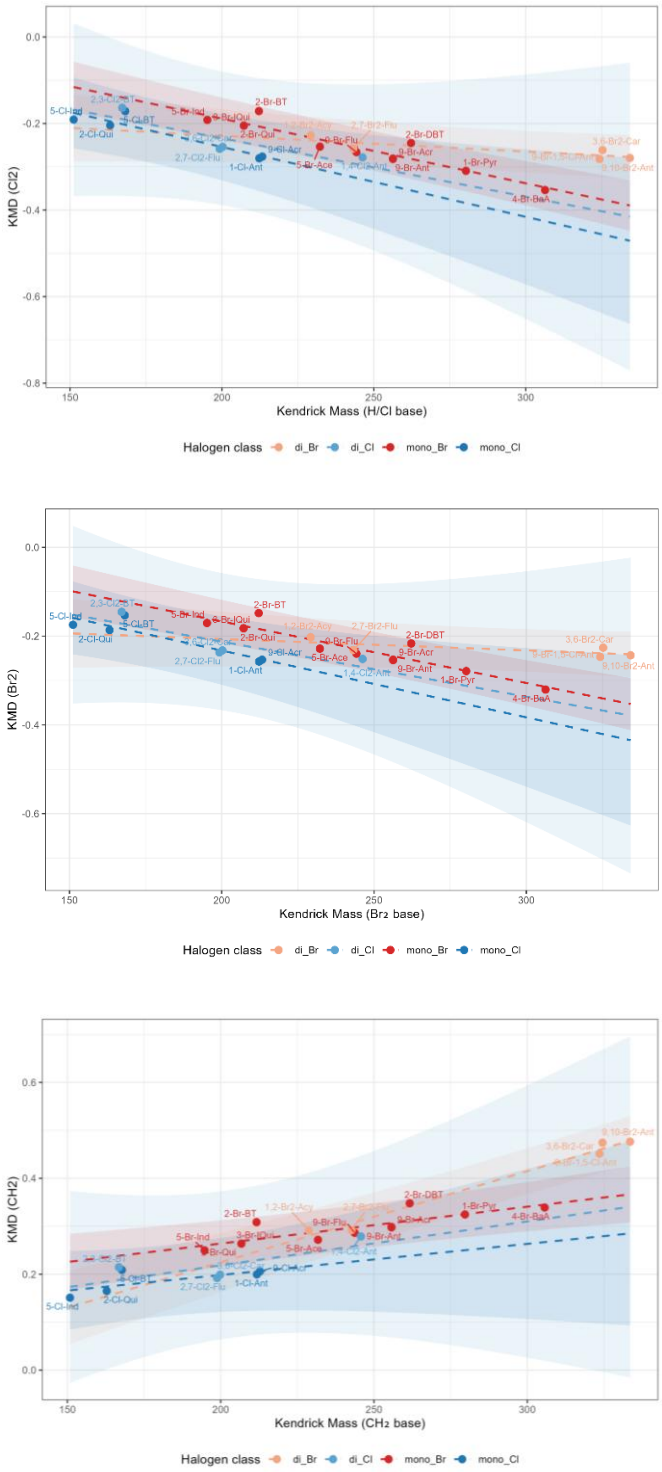



**Figure 8.3:** Kendrick mass defect plots created for halogenated PACs standards using H/Cl (top panel), $Br_2$ (middle panel), and $CH_2$ (bottom panel). Colored bands show 95% predication intervals for each compound class.

### 8.4.5. Workflow 3: identification of potential heterocyclic aromatic compounds

Similar to workflow 2, initial filtering steps were performed to remove low rt1 and m/z values, as well as column bleed peaks, thereby reducing background interference and improving the reliability of downstream screening. An evidence-based scoring system was then applied to identify compounds with a high likelihood of being HPACs.

The first scoring criterion involved comparison of mass fragment ions of peaks selected in the previous filtering steps against a reference database for PACs (Table V 10). Because the base peak does not necessarily correspond to the molecular ion, all detected fragment masses were compared to the reference exact masses, and fragments with at least one match within ±5 ppm were retained.

Kendrick mass defect analysis, using the $CH_2$ as the base unit, was subsequently applied as an annotation tool rather than a strict filter. This approach facilitates the identification of homologous series differing by $CH_2$ units, which is characteristic of alkylated PACs. Peaks with KMD values > 0.05 were flagged as potentially heteroatom-containing compounds, providing additional structural evidence.

An isotopic verification step was then performed to assess the presence of M+1 and M+2 peaks relative to the base mass. The M+1 peak indicates the presence of $^{13}C$ isotopes, while M+2 may indicate contributions from heavier isotopes such as $^{34}S$. This step provides further supporting evidence of elemental composition and helps distinguish true compounds from background noise.

Next, a double bond equivalent (DBE) estimation was applied using a simplified criterion (mass/26 ≥ 4) to identify fragments consistent with aromatic systems. Given that PACs are inherently aromatic, this step excludes non-aromatic and highly saturated compounds. Additional evidence was obtained through neutral loss analysis, in which all pairwise mass differences (±5 ppm) were

evaluated to identify characteristic losses, such as HCN (N-containing), CO (O-containing), and S or $H_2S$ (S-containing).

All criteria were integrated into a quantitative scoring system, yielding a maximum evidence score of eight. Scoring factors include KMD deviation, isotope presence, sulfur-specific isotope patterns, diagnostic fragment losses, and aromaticity. Compounds with scores > 5 were classified as HPACs. Finally, library-based annotation was performed by comparison with the PAC reference library created in WF1. Matching criteria included base mass (± 5 ppm), fragment ion overlap (minimum of two shared ions), and retention time alignment.

Application of this workflow to analytical standards resulted in 37 compounds being classified as possible HPACs, including three false positives (2-methylnapthalene, fluorene-d10, and acenaphthene-d10), while 15 HPACs were not identified. When applied to a yellow perch sample, the workflow reduced 15,000 initial peaks to 325 candidate HPACs.

The RECETOX Exposome library was applied to WF3 without modification to provide an external empirical evaluation of the HPAC screening logic. The library contains 23 HPACs among 386 spectra. At the reported high-confidence tier (evidence score ≥ 5), WF3 recovered 21 of the 23 target compounds (sensitivity 0.913; specificity 0.251; precision 0.072). Lowering the threshold to an evidence score of 3 recovered all 23 targets (Table V 11). WF3 therefore functions as a high-recall screen that does not discard heterocyclic PACs, but which returns a candidate list in which the target class remains a minority.

Examination of the individual scoring criteria revealed that the limited specificity stems from the composition of the evidence score rather than the performance of any single criterion. Of the eight possible points, five are assigned by criteria triggered for more than 90% of all detected features,

providing little discriminatory power for classifying heterocyclic PACs. These include the Kendrick mass defect criterion (double-weighted; triggered for 93.9% of features), the aromatic DBE criterion (99.2%), and the M+1 isotope criterion, which detects the naturally abundant $^{13}C$ envelope present in all carbon-containing compound. In contrast, only the diagnostic neutral-loss criteria demonstrated meaningful discrimination between heterocyclic PACs and background features, contributing just two of the eight points. Consequently, background compounds lacking any heteroatom-specific evidence still accumulate five points by default, explaining why 293 of 386 detected features were classified with the high-confidence tier.

This contrasts with the halogen filter in WF2, which achieved a sensitivity of 0.955 and specificity of 0.841 on the same library. The difference reflects the fundamentally different nature of the available analytical evidence rather than differences in implementation. The M+2 isotope pattern characteristic of $^{37}Cl$ and $^{81}Br$ is a feature that is shared by very few non-halogenated compounds. In contrast, nitrogen, sulfur, and oxygen provide no comparable diagnostic signature under electron ionization conditions. Their contributions to mass defect are small relative to carbon and hydrogen, while their characteristic neutral losses (HCN, CO, $H_2S$) are also produced by a wide range of non-aromatic and non-heterocyclic organic compounds. Consequently, heteroatom screening in EI spectra relies on the combined weight of several individually weak indicators, which, even when combined, cannot achieve the level of specificity afforded by halogen isotope patterns (Table V 12).

This asymmetry supports the organisation of the framework as four complementary workflows rather than a set of independent classifiers. WF2 is designed to make highly specific decisions within its scope, whereas WF3 prioritises sensitivity to maximise recovery of potential HPACs

during the initial screening stage. This strategy reflects the principle that a true HPAC excluded during screening cannot be recovered later in the workflow, whereas false-positive candidates can be eliminated during subsequent confirmation steps. Consequently, the low precision observed for WF3 is an expected characteristic of a first-pass screening workflow rather than a limitation of the overall framework.

Importantly, WF3 outputs are not regarded as compound identifications at any stage of this study. Final confirmation is achieved in WF4 using evidence that is independent of the spectral criteria employed in WF3, including two-dimensional chromatographic retention behaviour, in-house library spectral matching, and homologous series relationships identified through Kendrick mass defect analysis. Reported compound identifications therefore require concordance among multiple independent lines of evidence rather than reliance on the WF3 evidence score alone.

One important limitation of this benchmark should be acknowledged. The RECETOX library does not include chromatographic retention data, preventing evaluation of the retention-time-dependent confirmation tiers incorporated into the WF3 library annotation step. Consequently, only two of the 363 features received library confidence assignments, and the reported performance reflects only the spectral evidence component of WF3, without its chromatographic confirmation stage. The results should therefore be interpreted as a conservative estimate of the workflow's performance, with higher specificity expected when chromatographic retention information is available.

### 8.4.6. Workflow 4: results integration and PAC identification

This final workflow combines the results of the previous three workflows: library matching, halogenated PAC screening, and non-targeted HPACs detection to prioritize peaks with the highest

likelihood of being PACs. Final candidate selection was based on an evidence-based confidence scoring system.

A weighted score system was implemented to assign a final reliability score (0–10) to each peak. Library-based identifications were treated as the dominant source of confidence due to the robustness of retention time alignment and mass spectral similarity. As such, peaks classified as Level 1-3 in WF1 were assigned high base scores, ensuring that well-characterized homocyclic PACs and alkylated PAHs remained prioritized, even in the absence of heteroatom- and/or halogen-specific evidence.

Additional evidence from the heteroatom (NSO) and halogen workflows was incorporated as supplementary scoring and served as a verification layer, enhancing our confidence in candidate assignments. This approach allows for the inclusion of novel or previously uncharacterized compounds that may lack reference mass spectra, a common limitation in non-targeted environmental analysis. To minimize redundancy and over-reporting, duplicate entries were removed by retaining a single representative entry with the highest reliability score. Peaks with final scores ≥ 5 were classified as possible PACs.

### 8.4.7. Targeted and non-targeted PACs were detected in the environmental samples

Three separate replicate GC injections of sediment (n=2) and pooled biota (n=2) samples, along with method blanks, were analyzed in random order to minimize analytical bias. Due to the higher detection limits of our GCxGC-HRTOF-MS relative to GC-MS/MS, individual extracts of 5 fish were pooled (n=1) and concentrated to 50 uL prior to GC-injection. Semi-quantitation of non-target analytes were determined using the response factor of the nearest eluting authentic standard.

Estimated concentrations are reported in ng/g and summarized in Table V 14. Compounds were classified based on workflow-derived confidence levels and highlighted in Table V 13.

Compared with previously published targeted GC-MS/MS results for yellow perch, the non-targeted GC×GC-HRTOF analysis reveals a much broader and more chemically diverse PAC profile (see KMD plots in Figure V 1). The targeted method quantified 54 PACs, including elevated concentrations of polyaromatic nitrogen heterocycles (PANHs) such as indole, halogenated indoles, quinolines, and several high-molecular-weight PAHs (e.g., benzo[a]pyrene and indeno[1,2,3-c,d]pyrene). These compounds were also detected in the NTA dataset.

The NTA workflow uncovered numerous compounds do not present in our targeted list, including highly alkylated species (e.g., 1,4,6,7-tetramethylnaphthalene and 2,3,5-trimethylnaphthalene), additional heterocyclics (e.g., methylated isoquinolines and carbazoles), and several halogenated PACs such as dichloronaphthalene ($Cl_2$-Nap), dibromodibenzofuran ($Br_2$-DBF), and trichlorodibenzofuran ($Cl_3$-DBF). PANHs were the most abundant compound class detected in targeted analysis, and NTA illustrates that PANHs levels are likely being underestimating. For instance, targeted GC-MS/MS analysis indicated elevated concentrations of 5-bromoindole, whereas NTA revealed multiple isomers of this compound, highlighting limitations in compound coverage in targeted workflows.

Compared with the targeted GC-MS/MS analysis, which detected relatively few PACs in white bass and was dominated by a limited number of compounds, including indole, 5-bromoindole, C2-chrysene, and alkylated naphthalenes and fluorenes, the non-targeted GC×GC-HRTOF-MS workflow revealed a substantially broader and more chemically diverse PAC profile. Major compounds such as indole, phenanthrene, fluorene, carbazole, and quinoline were consistently

detected by both approaches, demonstrating good agreement for the dominant constituents. However, several abundant targeted analytes, including C2-chrysene, alkylated fluorenes, and halogenated quinolines, were not identified by the non-targeted workflow, likely reflecting a combination of lower analytical sensitivity and limitations associated with automated spectral deconvolution and library matching. Conversely, the non-targeted analysis identified numerous compounds absent from the targeted method, including alkylated homologues (e.g., 1,4-dimethylnaphthalene and 2,3,5-trimethylnaphthalene), structurally diverse heterocyclic PACs, halogenated PACs (e.g., tetrachloronaphthalene, tribromodibenzofuran, and dichlorocarbazole isomers), and multiple structural isomers that could not be resolved by the targeted GC-MS/MS analysis.

These findings demonstrate that, while targeted analysis provides highly sensitive and quantitative measurements for established priority compounds, it captures only a limited fraction of the chemical complexity present in environmental mixtures. In contrast, NTA expands chemical coverage by revealing a broader range of homologous series, transformation products, and previously unmonitored compounds, although it may be less effective for detecting low-abundance included in targeted methods. Together, these results highlight the complementary roles of targeted and non-targeted approaches, with NTA providing an important capability for discovering emerging and potentially environmentally relevant contaminants beyond conventional monitoring frameworks.

The GC-MS/MS analysis revealed that Hamilton Harbour sediment samples contained more than 90% of the targeted analytes in detectable quantities. The most abundant compounds included abundant parent PAHs (e.g., fluoranthene, pyrene, chrysene, benzo[g,h,i]perylene), many APAHs

(e.g., C2–C4 naphthalenes, phenanthrenes, and fluorenes), and a substantial number of PANHs (e.g., indole, carbazole, benzocarbazoles). Therefore, those samples were directly injected into HRTOF without a further concentration step.

The sediment sample exhibited high overall concentrations of parent and alkylated PAHs, and like biota, these sediments showed strong enrichment of heterocyclic compounds. In both cases, dominant compounds such as naphthalene, phenanthrene, fluorene, carbazole, and indole were consistently detected by the non-targeted GC×GC-HRTOF analysis, demonstrating good agreement between the methods at high analyte concentrations. However, the NTA workflow substantially expands the detected chemical space in sediment samples by revealing numerous additional compounds not included in the targeted list, including multiple isomers of highly alkylated PACs (e.g., dimethyl-, trimethyl-, and tetramethylnaphthalenes and dibenzothiophenes), a wider range of heterocyclic species, and halogenated PACs (e.g., halogenated anthracenes, chlorinated carbazoles, and halogenated fluorenes). Additionally, some targeted compounds were not always clearly resolved in the NTA data due to co-elution and spectral complexity at elevated concentrations. Overall, while targeted analysis effectively quantifies the magnitude and major contributors of PAC contamination in sediments, the NTA results demonstrate that both samples contain more structurally complex compounds, emphasizing the importance of non-targeted approaches for capturing isomer diversity, substitution patterns, and previously unmonitored compounds.

### 8.4.8. Rationale for Developing a Novel Workflow

A growing number of automated tools have been developed to facilitate the discovery of non-targeted environmental contaminants from HRMS datasets.[31] However, most existing approaches

have been optimized for broad chemical annotation or specific contaminant classes and do not adequately address the unique challenges associated with comprehensive PAC identification. The present four-workflow framework was developed to overcome this analytical gap by enabling the non-targeted detection and prioritization of homocyclic PACs, heterocyclic PACs, and halogenated PACs in complex environmental matrices analyzed by GC×GC–HRTOF–MS. To date, no published workflow was identified that provides comparable coverage across these chemically diverse PAC classes while integrating complementary evidence streams required for confident annotation. PAH-Finder[20] represents the closest related published approach and was therefore considered during the development of the present framework.

To assess the applicability of existing automated approaches for PAC discovery, PAH-Finder was initially applied here without modification to authentic PAC and HPAC standards acquired by both GC×GC-HRTOF-MS and conventional GC-HRTOF-MS. PAH-Finder was selected as the most closely related published approach and defines suspect spectra based on random forest classification of fragment distributions combined with detection of doubly charged fragment ions. However, the model was developed using single-column GC-QTOF-MS data and trained primarily on PAHs and alkylated PAHs, without representation of the broader HPAC chemical space targeted in this study.

Although doubly charged fragments provide useful structural information in EI spectra, their corresponding $^{13}C$ isotope peaks occur at low abundance (approximately 1–4% of the base peak intensity in the present datasets) and were frequently removed by our filtering criteria. When applied directly to authentic standards, PAH-Finder classified only approximately 9–10% of PAC-like compounds as suspects on both platforms, indicating that the doubly charged fragment

requirement represented the primary limiting factor under the acquisition conditions evaluated. These results should therefore be interpreted as an assessment of method transferability rather than a re-evaluation of the original workflow. They highlight the need for alternative evidence streams that remain consistently observable in EI GC-HRMS data, such as exact mass, halogen isotope patterns, and homologous-series relationships.

Other available computational tools were not directly applicable to our analytical challenge. ChloroDBPFinder was developed for chlorinated disinfection by-products using LC-HRMS electrospray ionization data and does not represent the broader PAC chemical space investigated here. Similarly, FIDDLE, a deep-learning approach for molecular formula prediction from LC-MS/MS data, has limited performance for halogenated compounds due to insufficient representation of these structures in its training datasets.

Together, these limitations demonstrate the need for a workflow specifically designed for PAC discovery using EI GC×GC-HRTOF-MS. The present framework was therefore developed around multiple independent evidence streams that are chemically informative and experimentally observable under these conditions. Rather than relying on a single diagnostic criterion or sequential gating strategy, WF4 integrates complementary evidence into an overall confidence score, enabling prioritization of PAHs, APAHs, HPAHs, and HPACs in complex environmental matrices where no dedicated identification workflow currently exists.

## 8.5 CHALLENGES AND LIMITATIONS OF THE DATA ANALYSIS PROCEDURE

Several limitations were identified during workflow application. WF1 relies on the coverage of the in-house spectral library, and compounds outside this chemical space receive only preliminary NIST-based annotation. Current library coverage remains incomplete for structurally diverse

PACs; for example, a single APAH isomer may represent an entire homologous class (e.g., 9-butylfluorene as a surrogate for C4-fluorenes), limiting confidence in isomer-level assignments. In addition, retention-time-based identifications require periodic library updates following column maintenance, while closely eluting compounds such as naphthalene derivatives may be challenging to resolve due to subtle second-dimension retention differences.

Analytical limitations also influence detection capability. Modulation conditions can affect peak shape and resolution, while the lower sensitivity of GC×GC-HRTOF-MS compared with targeted tandem MS approaches restricts detection of low-abundance compounds, particularly halogenated PACs in sediment matrices. Semi-quantitative estimates introduce further uncertainty because concentrations are calculated using the response of the nearest eluting targeted surrogate, despite potentially large differences in ionization efficiency among PAC classes.

Finally, automated annotation remains challenged by structurally similar compounds, as demonstrated by false assignments of PCBs as halogenated acenaphthene-like compounds. Although the workflow substantially reduces data complexity, expert review is still required for final confirmation of ambiguous features and validation of high-confidence identifications.

## 8.6 FUTURE DIRECTIONS

Future improvements should focus on expanding chemical coverage, increasing automation, and strengthening compound confirmation. Expansion of the in-house PAC library to include multi-substituted halogenated PACs and additional APAH isomers will improve WF1 coverage and reduce reliance on lower-confidence annotations. Development of a fully automated end-to-end pipeline incorporating machine-learning-based peak confirmation and identity resolution would further minimize manual review beyond WF4. Advanced classification approaches, including

random forests and neural networks trained on high-resolution mass spectral features, could enhance both sensitivity and specificity. Automated retention-time correction algorithms would also reduce the burden of library maintenance by accounting for column-to-column and day-to-day chromatographic variation. Finally, incorporation of negative chemical ionization, particularly electron-capture-based ion sources, could substantially improve the detection of halogenated PACs while generating simplified high-resolution spectra that provides an additional complementary line of evidence for confident identification.

## 8.7 ACKNOWLEDGMENTS

This work was supported by a Natural Sciences and Engineering Research Council Discovery (NSERC) Grant to GTT (Grant number: RGPIN-2024-04947). NV was supported by a University of Manitoba Graduate Fellowship, BP received an NSERC summer undergraduate award, and TT received a University of Manitoba Research Assistant award.

## 8.8 REFERENCES


((1) Marvin, C. H.; Berthiaume, A.; Burniston, D. A.; Chibwe, L.; Dove, A.; Evans, M.; Hewitt, L. M.; Hodson, P. V.; Muir, D. C. G.; Parrott, J.; Thomas, P. J.; Tomy, G. T. Polycyclic Aromatic Compounds in the Canadian Environment: Aquatic and Terrestrial Environments. *Environmental Pollution*. Elsevier Ltd September 15, 2021. https://doi.org/10.1016/j.envpol.2021.117442.

(2) Berthiaume, A.; Galarneau, E.; Marson, G. Polycyclic Aromatic Compounds (PACs) in the Canadian Environment: Sources and Emissions. *Environmental Pollution*. Elsevier Ltd January 15, 2021. https://doi.org/10.1016/j.envpol.2020.116008.

(3) Marvin, C. H.; Tomy, G. T.; Thomas, P. J.; Holloway, A. C.; Sandau, C. D.; Idowu, I.; Xia, Z. Considerations for Prioritization of Polycyclic Aromatic Compounds as Environmental Contaminants. *Environmental Science and Technology*. American Chemical Society December 1, 2020, pp 14787–14789. https://doi.org/10.1021/acs.est.0c04892.

(4) Andersson, J. T.; Achten, C. Time to Say Goodbye to the 16 EPA PAHs? Toward an Up-to-Date Use of PACs for Environmental Purposes. *Polycycl. Aromat. Compd.* **2015**, *35* (2–4), 330–354. https://doi.org/10.1080/10406638.2014.991042.

(5) Machado, M. E.; Nascimento, M. M.; Bomfim Bahia, P. V.; Martinez, S. T.; Bittencourt de Andrade, J. Analytical Advances and Challenges for the Determination of Heterocyclic Aromatic Compounds (NSO-HET) in Sediment: A Review. *TrAC - Trends in Analytical Chemistry*. Elsevier B.V. May 1, 2022. https://doi.org/10.1016/j.trac.2022.116586.

(6) Vitharana, N. N.; Halldorson, T.; Xia, Z.; Francisco, O.; Marvin, C.; Thomas, P. J.; Liebzeit, E.; Lucas, A. M.; Moradi, V.; Tomy, G. T. A Validated Approach for Analysis of Heterocyclic Aromatic Compounds in Sediment Samples. *J. Chromatogr. A* **2024**, *1718*. https://doi.org/10.1016/j.chroma.2024.464723.

(7) Wang, Z. *Oil Spill Environmental Forensics*; Academic Press, 2007. https://doi.org/https://doi.org/10.1016/B978-0-12-369523-9.X5000-9.

(8) Johnson, W.; Idowu, I.; Francisco, O.; Marvin, C.; Thomas, P. J.; Stetefeld, J.; Tomy, G. T. Enumeration of the Constitutional Isomers of Environmentally Relevant Substituted Polycyclic Aromatic Compounds. *Chemosphere* **2018**, *202*, 9–16. https://doi.org/10.1016/j.chemosphere.2018.03.035.

(9) Idowu, I.; Francisco, O.; Thomas, P. J.; Johnson, W.; Marvin, C.; Stetefeld, J.; Tomy, G. T. Validation of a Simultaneous Method for Determining Polycyclic Aromatic Compounds and Alkylated Isomers in Biota. *Rapid Communications in Mass Spectrometry* **2018**, *32* (3), 277–287. https://doi.org/10.1002/rcm.8035.

(10) Vitharana, N. N.; Halldorson, T.; Paddock, B.; Tomy, T.; Marvin, C.; Fisk, A. T.; Heuvel, C. E.; Tomy, G. T. Habitat and Diet Affect Bioaccumulation of Homocyclic and Heterocyclic Aromatic

Compounds in Lake Erie Fish. *J. Hazard. Mater.* **2026**, *509*, 142044. https://doi.org/10.1016/j.jhazmat.2026.142044.

(11) Vitharana, N. N.; Halldorson, T.; Dereviankin, M.; Paddock, B.; Gerring, A.; Marvin, C.; Helm, P.; Thomas, P. J.; Idowu, I. G.; Tomy, G. T. Heterocyclic Aromatic Compounds in Lake Ontario Sediments: Spatiotemporal Distribution and New Approaches to Fingerprinting. *Science of the Total Environment* **2025**, *986*. https://doi.org/10.1016/j.scitotenv.2025.179802.

(12) Manzano, C. A.; Marvin, C.; Muir, D.; Harner, T.; Martin, J.; Zhang, Y. Heterocyclic Aromatics in Petroleum Coke, Snow, Lake Sediments, and Air Samples from the Athabasca Oil Sands Region. *Environ. Sci. Technol.* **2017**, *51* (10), 5445–5453. https://doi.org/10.1021/acs.est.7b01345.

(13) Xia, Z.; Idowu, I.; Marvin, C.; Thomas, P. J.; Johnson, W.; Francisco, O.; Stetefeld, J.; Crimmins, B.; Fry, M.; Tomy, G. T. Identification of Halogenated Polycyclic Aromatic Hydrocarbons in Biological Samples from Alberta Oil-Sands Region. *Chemosphere* **2019**, *215*, 206–213. https://doi.org/10.1016/j.chemosphere.2018.10.050.

(14) Liebzeit, E. C.; Vitharana, N. N.; Halldorson, T.; Provencher, J. F.; Mallory, M. L.; Tomy, G. T.; Dupuis-Smith, R. Species- and Tissue-Specific Accumulation of Polycyclic Aromatic Compounds in Three Arctic Seabirds. *Environ. Sci. Technol.* **2026**. https://doi.org/10.1021/acs.est.5c09333.

(15) Canchola, A.; Tran, L. N.; Woo, W.; Tian, L.; Lin, Y. H.; Chou, W. C. Advancing Non-Target Analysis of Emerging Environmental Contaminants with Machine Learning: Current Status and Future Implications. *Environment International*. Elsevier Ltd April 1, 2025. https://doi.org/10.1016/j.envint.2025.109404.

(16) Chibwe, L.; Manzano, C. A.; Muir, D.; Atkinson, B.; Kirk, J. L.; Marvin, C. H.; Wang, X.; Teixeira, C.; Shang, D.; Harner, T.; De Silva, A. O. Deposition and Source Identification of Nitrogen Heterocyclic Polycyclic Aromatic Compounds in Snow, Sediment, and Air Samples from the Athabasca Oil Sands Region. *Environ. Sci. Technol.* **2019**, *53* (6), 2981–2989. https://doi.org/10.1021/acs.est.8b06175.

(17) Fernando, S.; Renaguli, A.; Milligan, M. S.; Pagano, J. J.; Hopke, P. K.; Holsen, T. M.; Crimmins, B. S. Comprehensive Analysis of the Great Lakes Top Predator Fish for Novel Halogenated Organic Contaminants by GC×GC-HR-ToF Mass Spectrometry. *Environ. Sci. Technol.* **2018**, *52* (5), 2909–2917. https://doi.org/10.1021/acs.est.7b05999.

(18) Achten, C.; Marin-Enriquez, O.; Behrends, B.; Kupich, S.; Lutter, A.; Korth, R.; Andersson, J. T. Polycyclic Aromatic Compounds Including Non-Target and 71 Target Polycyclic Aromatic Hydrocarbons in Scrubber Discharge Water and Their Environmental Impact. *Mar. Pollut. Bull.* **2024**, *208*. https://doi.org/10.1016/j.marpolbul.2024.116790.

(19) Yang, L.; Wu, J.; Zheng, M.; Cao, Z.; Li, C.; Shi, M.; Liu, G. Non-Target Screening of Organic Pollutants and Target Analysis of Halogenated Polycyclic Aromatic Hydrocarbons in the Atmosphere around Metallurgical Plants by High-Resolution GC/Q-TOF-MS. *Environ. Sci. Eur.* **2020**, *32* (1). https://doi.org/10.1186/s12302-020-00376-9.

(20) Zhang, Z.; Xu, X.; Xing, S.; Shi, C.; You, Z.; Deng, X.; Tan, L.; Mo, Z.; Fang, M. PAH-Finder: A Pattern Recognition Workflow for Identification of PAHs and Their Derivatives. *Anal. Chem.* **2025**, *97* (2), 1170–1179. https://doi.org/10.1021/acs.analchem.4c04249.

(21) Zhao, T.; Wawryk, N. J. P.; Xing, S.; Low, B.; Li, G.; Yu, H.; Wang, Y.; Shen, Q.; Li, X. F.; Huan, T. ChloroDBPFinder: Machine Learning-Guided Recognition of Chlorinated Disinfection Byproducts from Nontargeted LC-HRMS Analysis. *Anal. Chem.* **2024**, *96* (6), 2590–2598. https://doi.org/10.1021/acs.analchem.3c05124.

(22) Hong, Y.; Li, S.; Ye, Y.; Tang, H. FIDDLE: A Deep Learning Method for Chemical Formulas Prediction from Tandem Mass Spectra. *Nature Communications* **2025**, *16* (1). https://doi.org/10.1038/s41467-025-66060-9.

(23) Idowu, I. G.; Dereviankin, M.; Halldorson, T.; Vitharana, N.; Harbicht, L.; Helm, P.; Kleywegt, S.; Marvin, C.; Sandau, C. D.; Tomy, G. T. Source Characterization of Polycyclic Aromatic Hydrocarbons in Urban Sediments: Insights from the Contribution of Coal Tar, Sealants, and Other Related Industrial Materials in Lake Ontario. *Science of the Total Environment* **2025**, *1009*. https://doi.org/10.1016/j.scitotenv.2025.180967.

(24) Vitharana, N. N.; Halldorson, T.; Paddock, B.; Tomy, T.; Gerring, A.; Marvin, C.; Heuvel, C. E.; Fisk, A. T.; Tomy, G. T. Using Chemometrics to Discern Patterns and Spatial Distribution of Heterocyclic Aromatic Compounds in Lake Erie Fish. *Environmental Pollution* **2026**, *399*, 128121. https://doi.org/10.1016/j.envpol.2026.128121.

(25) *The Fitness for Purpose of Analytical Methods : A Laboratory Guide to Method Validation and Related Topics*, Second edition, 2014.; Magnusson, B., Örnemark, U., Eds.

(26) Idowu, I.; Johnson, W.; Francisco, O.; Obal, T.; Marvin, C.; Thomas, P. J.; Sandau, C. D.; Stetefeld, J.; Tomy, G. T. Comprehensive Two-Dimensional Gas Chromatography High-Resolution Mass Spectrometry for the Analysis of Substituted and Unsubstituted Polycyclic Aromatic Compounds in Environmental Samples. *J. Chromatogr. A* **2018**, *1579*, 106–114. https://doi.org/10.1016/j.chroma.2018.10.030.

(27) Official Journal of the European Communities (2002) Commission Decision of 12 August 2002 Implementing Council Directive 96/23/EC Concerning the Performance of Analytical Methods and the Interpretation of Results.; 2002.

(28) Jobst, K. J.; Shen, L.; Reiner, E. J.; Taguchi, V. Y.; Helm, P. A.; McCrindle, R.; Backus, S. The Use of Mass Defect Plots for the Identification of (Novel) Halogenated Contaminants in the Environment. *Anal. Bioanal. Chem.* **2013**, *405* (10), 3289–3297. https://doi.org/10.1007/s00216-013-6735-2.

(29) ’Price, E. J. ’, ’Palát, J. ’, ’Coufaliková, K. ’, ’Kukučka, P. ’, ’Codling, G. , ’Vitale, C. M. ’, K. Š. , ’Klánová, J. ’. RECETOX Exposome HR-[EI+]-MS Library (Version 1.0.0) [Dataset]. Zenodo. **2020**.

(30) Merel, S. Critical Assessment of the Kendrick Mass Defect Analysis as an Innovative Approach to Process High Resolution Mass Spectrometry Data for Environmental Applications. *Chemosphere*. Elsevier Ltd February 1, 2023. https://doi.org/10.1016/j.chemosphere.2022.137443.

(31) Wang, H.; Zhong, L.; Su, W.; Ruan, T.; Jiang, G. Machine Learning-Assisted Identification of Environmental Pollutants by Liquid Chromatography Coupled with High-Resolution Mass Spectrometry. *TrAC - Trends in Analytical Chemistry*. Elsevier B.V. November 1, 2024. https://doi.org/10.1016/j.trac.2024.117988.

# CHAPTER 9: CONCLUSIONS AND FUTURE DIRECTIONS

## 9.1 CONCLUSIONS

My PhD research aimed to develop analytical procedures to extract, purify, and detect hetero-polycyclic aromatic compounds and to identify their sources, fate, and behavior using advanced chemometric techniques. Based on the heteroatom type, alkylation, and halogenation, HPACs show a wide range of physical and chemical properties. Therefore, simultaneously extracting all HPAC types along with homocyclic PACs is difficult. Also, structural complexity limits the availability of authentic analytical standards, hindering the development of targeted analytical procedures capable of detecting and quantifying a broad range of HPACs.

In this study, I have generated new knowledge that advances the field of environmental chemistry. Specifically, I developed analytical methods for simultaneous extraction, purification, and detection of all three types of HPACs; generated baseline environmental concentration data for Lake Ontario and Lake Erie for the first time for numerous PACs; demonstrated that HPACs to be superior chemical fingerprints than the homocyclic PACs; established bioaccumulation patterns of HPACs and the environmental and physiological factors that contribute to that and, finally, created an advanced data filtering workflow to screen PACs from complex high-resolution mass spectral data.

I developed targeted analytical methods using GC-MS/MS to detect 50 HPACs. Extraction and purification procedures were developed for the simultaneous analysis of nitrogen-, sulfur-, and/or oxygen-containing HPACs in sediment and fish muscle tissue samples. All developed analytical procedures were validated in accordance with Eurachem guidelines at an ISO 17025-accredited

laboratory. The performance characteristics indicated that the developed methods were fit for the intended purpose.

Using the developed and validated targeted analytical methods, I analysed sediments from Lake Ontario and ten fish species from Lake Erie to identify the environmental distribution and behaviour of HPACs relative to homocyclic PACs. Overall, 104 targeted analytes were studied in these samples. The concentration data indicated that many HPACs were detectable in both biotic and abiotic samples. In most cases, PANHs were the most abundant HPAC type detected, and my data now serve as baseline data for future environmental risk assessment studies in the Great Lakes region.

In Lake Ontario sediments, HPAC load varied based on the location, with the rank order of total HPAC concentrations being Hamilton Harbour > Toronto Harbour > non-harbor areas. The HPAC profiles for the sediment samples and 14 different reference materials were compared, and multivariate analysis indicated that harbour sediments are more likely to be contaminated with coal-tar-based products. My results demonstrated that HPACs can serve as effective fingerprints for source apportionment studies of PACs. Furthermore, my study highlights the importance of integrating novel fingerprinting approaches, such as UMAP, with traditional multivariate techniques like PCA to achieve more comprehensive source characterization of complex environmental pollutant mixtures.

The concentration profiles of targeted PACs in four fish species (freshwater drum, yellow perch, gizzard shad, and walleye) from three basins (Western, Central, and Eastern) of Lake Erie indicated both species- and location-dependent variations. To my knowledge, this is the first targeted analytical study carried out to detect all three classes of HPACs in the Great Lakes biota. Also, this

is one of the first targeted analytical studies done to detect this many HPACs in any biota samples. My results demonstrated that the Western basin is the most contaminated basin of the Lake, with the highest concentrations across all PAC subclasses. Walleye samples exhibited the highest variability in HPAC profiles among the four fish species.

Ten fish species from the western basin of Lake Erie were also analysed for PACs. Alkylated PAHs and PANHs dominated PAC contamination in all fish species. Redundancy analysis indicated HPAC profiles varied notably among species and were strongly influenced by habitat and feeding ecology, whereas homocyclic PACs reflected uniform contamination patterns. Limited biomagnification was observed, with most analytes showing non-significant or low trophic magnification factors.

Multivariate and cluster analysis conducted individually on heterocyclic and homocyclic PAC profiles for biota and sediment samples indicated distinct profiles for HPACs and uniform profiles for PAHs and APAHs in most cases. In source fingerprinting studies, there was a clear separation between sediment samples from different locations and overlap between coal-tar-based reference materials and harbour sediments according to the multivariate analysis outputs. However, these clear clusters were not available when PAH and APAH concentrations in the same samples were compared.

Similarly, in Lake Erie fish, I observed distinct basin-, species-, and habitat-dependent clusters for HPACs. However, for homocyclic PACs, these subtle variations were not clearly identified using the same chemometric techniques. Taken together, my studies highlight that HPACs are better representatives of environmental occurrence and behaviour of PACs compared to PAHs and APAHs. This highlights the importance of incorporating HPACs into routine monitoring

frameworks, not only to identify the overall risk of PACs but also to effectively monitor their environmental movement.

One of the main challenges with non-targeted analytical procedures for organic pollutants is the volume of data generated for complex environmental matrices. The available spectral library data for PACs is insufficient to reliably identify them in environmental samples using non-targeted data. Therefore, in my study, non-targeted analysis using GCxGC-HRTOF-MS data was combined with advanced post-acquisition filtering workflows to improve the identification of different subclasses of PACs. Application of the filtering workflow to analytical standards and environmental samples from Lake Erie fish and Hamilton Harbour sediments successfully identified both targeted and previously unmonitored PAC classes. Overall, the workflows expanded detectable PAC chemical space while providing a reproducible and transferable strategy for environmental monitoring.

## 9.2 FUTURE DIRECTIONS

Throughout my study, key knowledge gaps and methodological limitations in the environmental analysis of HPACs were identified, highlighting several opportunities for future research and analytical advancement. One of the major challenges that needs to be addressed is the lack of standardized methods available for analysing underrepresented PAC subclasses in environmental matrices. This hinders the identification of the overall environmental impact of PACs.

Another challenge I faced during the method development phase was the lack of standard reference materials (SRM) for these compounds. There is a limited number of PAC compounds identified in commercially available SRMs. Therefore, developing an SRM that includes all PAC subclasses for both abiotic samples is important for implementing proper QA/QC measures in these studies.

Source fingerprinting studies rely on persistent environmental pollutants within the matrix under consideration. However, the environmental persistence data available for HPACs are limited. Especially for biota samples, persistence analysis studies needed to be carried out to take advantage of the superior fingerprinting power of HPACs.

Diagnostic ratio plots are traditionally constructed for PAHs and APAHs to identify petrogenic, pyrogenic, and biogenic sources. Despite several limitations associated with the technique, diagnostic ratios remain widely used in initial fingerprinting studies. My research highlights the greater source specificity of HPACs compared to homocyclic PACs. Hence, it is worthwhile to identify diagnostic ratios between common HPAC isomer pairs (e.g., acridine and benzo(h)quinoline) and between unsubstituted and alkylated HPAC pairs in environmental samples and compare them with reference materials.

To fully identify the sources of these compounds, source apportionment studies were needed. This would include chemical mass balancing approaches that go beyond the simple source fingerprinting approaches used in my study. This will ensure accurate identification of HPAC sources in these environmental systems.

To identify the bioconcentration and biomagnification factors of the detected compounds in biota, controlled laboratory studies are needed. Biotransformation studies will also increase the understanding of the impact of HPACs in food web studies. Apart from species-dependent characteristics, such as isotope ratios and morphology, incorporating environmental parameters (e.g., temperature, pH, and salinity) and HPAC concentrations in water into multivariate analysis will provide better insight into HPAC profiles in biota samples collected from complex environmental systems such as Lake Erie. Identification of transformation products of PACs, along

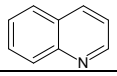
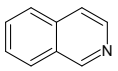
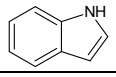
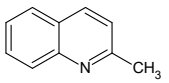
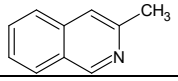
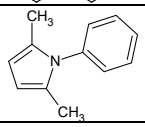
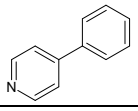
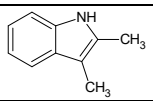
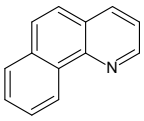
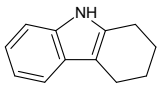
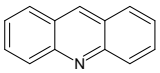
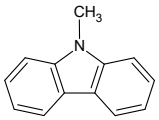
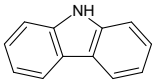
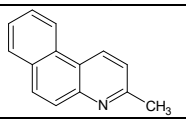
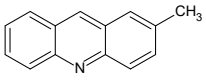
with their persistence, bioaccumulation, and toxicity, is also important for assessing the overall environmental impact of these pollutants.

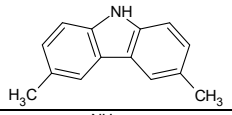
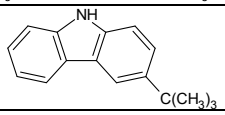
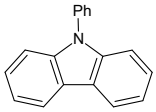
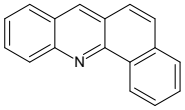
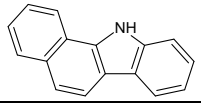
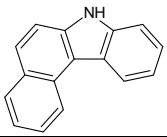
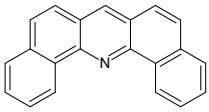
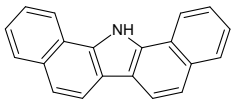
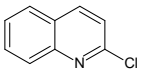
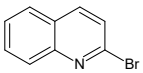
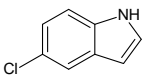
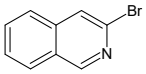
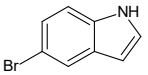
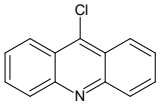
A major challenge I faced during non-targeted analysis of these compounds was the unavailability of the spectral data in commercial libraries. Updating these mass spectral reference libraries for HPACs and other underrepresented subclasses is recommended to improve the efficiency of non-targeted PAC analysis. Future improvements for my NTA data filtering workflow include expanding the in-house PAC library to cover multi-substituted halogenated PACs and APAH isomers to improve WF1 coverage and accuracy. Developing a fully automated pipeline with machine learning-based compound classification, automated peak confirmation, and retention time correction could enhance sensitivity and specificity while reducing manual processing and library maintenance. Additionally, using electron-capture negative ionization may improve the detection and identification of halo-PACs through simpler, high-resolution mass spectra.

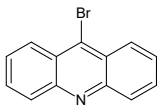
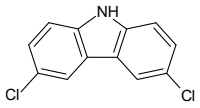
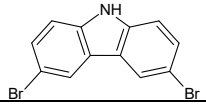
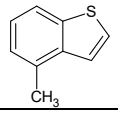
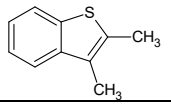
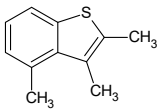
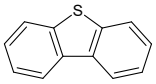
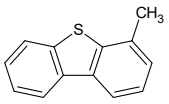
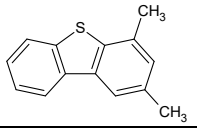
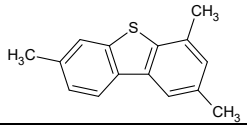
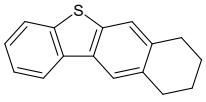
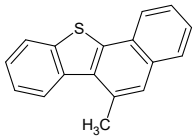
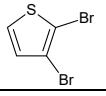
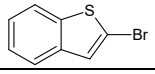
Although the mechanism of AhR activation by PAHs is well established, recent studies show that alkylation and halogenation can substantially alter both AhR activation potency and toxicity in fish models. However, the structure-activity relationships (SARs) between HPACs and the AhR, particularly for alkylated and halogenated analogues, remain poorly characterized. Future work should focus on integrating AhR-mediated in vitro toxicity assessments with molecular modeling approaches to better understand the SARs of HPAC derivatives. Such studies would improve understanding of how chemical substitutions influence toxicity and receptor interactions, ultimately leading to more accurate ecological risk assessments of complex PAC mixtures.

# APPENDIX I. ADDITIONAL METRICS FROM CHAPTER 3: A VALIDATED APPROACH FOR ANALYSIS OF HETEROCYCLIC AROMATIC COMPOUNDS IN SEDIMENT SAMPLES

**Table I 1:** General information regarding the standards included in the two MRM methods.

| CAS Number | The Chemical Name | Chemical Formula | Chemical Structure | Molecular Weight in g/mol | Boiling Point in °C ≈ |
|---|---|---|---|---|---|
| 91-22-5 | Quinoline | $C_9H_7N$ |  | 129.16 | 237.2 |
| 119-65-3 | Isoquinoline | $C_9H_7N$ |  | 129.16 | 243.2 |
| 120-72-9 | Indole | $C_8H_7N$ |  | 117.15 | 253 - 254 |
| 91-63-4 | 2-Methylquinoline | $C_{10}H_9N$ |  | 143.18 | 246 - 248 |
| 1125-80-0 | 3-Methylisoquinoline | $C_{10}H_9N$ |  | 143.18 | 251 |
| 83-24-9 | 2,5-Dimethyl-1-phenylpyrrole | $C_{12}H_{13}N$ |  | 171.24 | 155 - 160 |
| 939-23-1 | 4-Phenylpyridine | $C_{11}H_9N$ |  | 155.20 | 281 |
| 91-55-4 | 2,3-Dimethylindole | $C_{10}H_{11}N$ |  | 145.21 | 285 |
| 85-06-3 | Benzo[h]quinoline | $C_{13}H_9N$ |  | 179.22 | 338 |
| 942-01-8 | 1,2,3,4-Tetrahydrocarbazole | $C_{12}H_{13}N$ |  | 171.24 | 325-330 |
| 260-94-6 | Acridine | $C_{13}H_9N$ |  | 179.22 | 345.5 |
| 1484-12-4 | 9-Methylcarbazole | $C_{13}H_{11}N$ |  | 181.24 | 195 |
| 86-74-8 | Carbazole | $C_{12}H_9N$ |  | 167.21 | 354.6 |
| 1985-06-03 | 3-Methylbenzo[f]quinoline | $C_{14}H_{11}N$ |  | 193.24 | 352.5 |
| 613-15-0 | 2-Methylacridine | $C_{14}H_{11}N$ |  | 193.25 | - |

| | | | | | |
|---|---|---|---|---|---|
| 5599-50-8 | 3,6-Dimethylcarbazole | $C_{14}H_{13}N$ |  | 195.26 | - |
| 22401-74-7 | 3-Tert-butyl-9H-carbazole | $C_{16}H_{17}N$ |  | 223.31 | - |
| 1150-62-5 | 9-Phenylcarbazole | $C_{18}H_{13}$ |  | 243.3 | - |
| 225-51-4 | Benz[c]acridine | $C_{17}H_{11}N$ |  | 229.27 | - |
| 239-01-0 | 11H-Benzo[a]carbazole | $C_{16}H_{11}N$ |  | 217.26 | - |
| 205-25-4 | 7H-benzo[c]carbazole | $C_{16}H_{11}N$ |  | 217.26 | - |
| 224-53-3 | Dibenz[c,h]acridine | $C_{21}H_{13}N$ |  | 279.3 | - |
| 239-64-5 | 13(H)-Dibenzo[a,i]carbazole | $C_{20}H_{13}N$ |  | 267.3 | - |
| 612-62-4 | 2-Chloroquinoline | $C_9H_6ClN$ |  | 163.6 | - |
| 2005-43-8 | 2-Bromoquinoline | $C_9H_6BrN$ |  | 208.05 | - |
| 17422-32-1 | 5-Chloroindole | $C_8H_6ClN$ |  | 151.6 | - |
| 34784-02-6 | 3-Bromoisoquinoline | $C_9H_6BrN$ |  | 208.05 | - |
| 10075-50-0 | 5-Bromoindole | $C_8H_6BrN$ |  | 196.04 | - |
| 1207-69-8 | 9-Chloroacridine | $C_{13}H_8ClN$ |  | 213.66 | - |

| | | | | | |
|---|---|---|---|---|---|
| 4357-57-7 | 9-bromoacridine | $C_{13}H_8BrN$ |  | 258.11 | - |
| 5599-71-3 | 3,6-Dichlorocarbazole | $C_{12}H_7Cl_2N$ |  | 236.09 | - |
| 6825-20-3 | 3,6-Dibromocarbazole | $C_{12}H_7Br_2N$ |  | 325.00 | - |
| 14315-11-8 | 4-Methylbenzothiophene | $C_9H_8S$ |  | 148.23 | - |
| 4923-91-5 | 2,3-Dimethylbenzothiophene | $C_{10}H_{10}S$ |  | 162.25 | - |
| 70021-44-2 | 2,3,4-Trimethylbenzothiophene | $C_{11}H_{12}S$ |  | 176.28 | - |
| 132-65-0 | Dibenzothiophene | $C_{12}H_8S$ |  | 184.26 | 332.5 |
| 7372-88-5 | 4-Methyldibenzothiophene | $C_{13}H_{10}S$ |  | 198.29 | - |
| 31317-18-7 | 2,4-Dimethyldibenzothiophene | $C_{14}H_{12}S$ |  | 212.31 | - |
| 216983-03-8 | 2,4,7-Trimethyldibenzothiophene | $C_{15}H_{14}S$ |  | 226.3 | - |
| 39097-15-9 | 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | $C_{16}H_{14}S$ |  | 238.3 | - |
| 4567-51-5 | 6-Methylbenzo[b]naphtho[2,1-d]thiophene | $C_{17}H_{12}S$ |  | 248.3 | - |
| 3140-93-0 | 2,3-Dibromothiophene | $C_4H_2Br_2S$ |  | 241.93 | - |
| 5394-13-8 | 2-Bromobenzothiophene | $C_8H_5BrS$ |  | 213.10 | - |

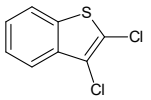
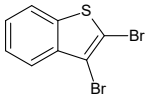
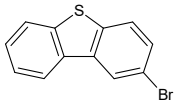
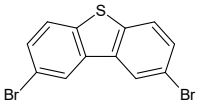
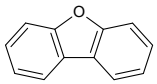
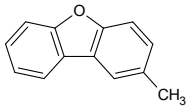
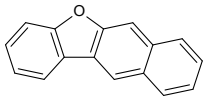
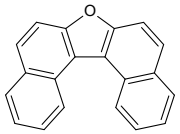
| | | | | | |
|---|---|---|---|---|---|
| 5323-97-7 | 2,3-Dichlorobenzothiophene | $C_8H_4Cl_2S$ |  | 203.09 | - |
| 6287-82-7 | 2,3-Dibromobenzothiophene | $C_8H_4Br_2S$ |  | 291.99 | - |
| 22439-61-8 | 2-Bromodibenzothiophene | $C_{12}H_7BrS$ |  | 263.15 | - |
| 31574-87-5 | 2,8-Dibromodibenzothiophene | $C_{12}H_6Br_2S$ |  | 342.05 | - |
| 132-64-9 | Dibenzofuran | $C_{12}H_8O$ |  | 168.19 | 287 |
| 7320-51-6 | 2-Methyldibenzofuran | $C_{13}H_{10}O$ |  | 182.22 | - |
| 243-42-5 | Benzo[b]naphtho[2,3-d]furan | $C_{16}H_{10}O$ |  | 218.25 | - |
| 194-63-8 | Dinaphtho[1,2-b;1',2'-d]furan | $C_{20}H_{12}O$ |  | 268.3 | - |

**Table I 2:** Percent recoveries of representative HPACs extracted (n=3) using DCM and DCM:ethyl acetate 1:1.

| Analyte | % Recovery with DCM | % Recovery with 1:1 DCM: ethyl acetate |
|---|---|---|
| Quinoline | 2 | 119 |
| Indole | 19 | 61 |
| 2-Methylquinoline | 3 | 105 |
| 3-Methylisoquinoline | 0 | 107 |
| 2,5-Dimethyl-1-phenylpyrrole* | 64 | 8 |
| 2,3-Dimethylindole* | 15 | 6 |
| Benzo[h]quinoline | 82 | 99 |
| 1,2,3,4-Tetrahydrocarbazole* | - | 32 |
| Acridine | 82 | 98 |
| 9-Methylcarbazole | 97 | 103 |
| Carbazole | 85 | 82 |
| 3-Methylbenzo[f]quinoline | 6 | 107 |
| 2-Methylacridine | 7 | 117 |
| 3,6-Dimethylcarbazole | 116 | 128 |
| 3-Tert-butyl-9H-carbazole | 118 | 118 |
| 9-Phenylcarbazole | 128 | 109 |
| Benz[c]acridine | 116 | 115 |
| 11H-Benzo[a]carbazole | 122 | 103 |
| 7H-benzo[c]carbazole | 116 | 92 |
| Dibenz[c,h]acridine | 183 | 87 |
| 13(H)-Dibenzo[a,i]carbazole | 196 | 89 |
| Dibenzothiophene | 101 | 91 |
| 4-Methyldibenzothiophene | 100 | 92 |
| 2,4-Dimethyldibenzothiophene | 152 | 111 |
| 2,4,7-Trimethyldibenzothiophene | 134 | 95 |
| 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | 165 | 100 |
| 6-Methylbenzo[b]naphtho[2,1-d]thiophene | 147 | 85 |
| Dibenzofuran | 99 | 116 |
| 2-Methyldibenzofuran | 98 | 98 |
| Benzo[b]naphtho[2,3-d]furan | 102 | 101 |
| Dinaphtho[1,2-b;1',2'-d]furan | 133 | 110 |

The analytes marked with * were removed from the method due to low recovery.

**Table I 3:** Mean percent recoveries (± standard deviation, n = 8) of RIS based on IPIS (anthracene-$d_{10}$) during the method validation

| ISTD | % Recovery of RIS at each concentration level | | | |
|---|---|---|---|---|
| | **C1** | **C2** | **C3** | **C4** |
| Naphthalene-$d_8$ | 83 ± 5.3 | 71 ± 9.2 | 83 ± 5.5 | 71 ± 8.4 |
| Fluorene-$d_{10}$ | 78 ± 4.3 | 77 ± 8.7 | 82 ± 6.8 | 88 ± 7.6 |
| Phenantherene-$d_{10}$ | 94 ± 3.5 | 91 ± 9.2 | 94 ± 5.1 | 97 ± 4.6 |
| Fluoranthene-$d_{10}$ | 86 ± 3.2 | 86 ± 8.4 | 92 ± 5.5 | 88 ± 14.8 |
| Pyrene-$d_{10}$ | 86 ± 3.2 | 86 ± 8.7 | 89 ± 4.8 | 93 ± 10.4 |
| Chrysene-$d_{12}$ | 99 ± 5.5 | 101 ± 6.6 | 94 ± 5.0 | 10 ± 10.9 |
| Benz[a]anthracene-$d_{12}$ | 105 ± 4.8 | 110 ± 14.6 | 100 ± 6.2 | 98 ± 13.7 |
| Benzo[b]fluranthene-$d_{12}$ | 101 ± 5.0 | 103 ± 10.0 | 87 ± 7.4 | 75 ± 12.4 |
| Benzo[ghi]perylene-$d_{14}$ | 95 ± 3.9 | 110 ± 10.2 | 88 ± 7.0 | 75 ± 10.7 |

'C1 = 800 pg/µL, C2 = 100 pg/µL, C3 = 25 pg/µL and C4 = 10 pg/µL' of HPACs

**Table I 4:** Mean percent recoveries (±SD) of native PAHs with the extraction method developed for HPACs.

| Standard Name | Percent Recovery |
|---|---|
| Acenapthene | 96 ± 2 |
| Acenapthylene | 91 ± 8 |
| Anthracene | 83 ± 10 |
| Benz[a]anthracene | 93 ± 2 |
| Benzo[a]pyrene | 86 ± 1 |
| Benzo[b]fluoranthene | 93 ± 1 |
| Benzo[g,h,i]perylene | 91 ± 2 |
| Benzo[k]fluoranthene | 117 ± 11 |
| Chrysene | 90 ± 2 |
| Dibenzo[a,h]anthracene | 97 ± 1 |
| Fluoranthene | 96 ± 3 |
| Fluorene | 92 ± 3 |
| Indeno[1,2,3-c,d]pyrene | 104 ± 2 |
| Naphthalene | 111 ± 2 |
| Phenanthrene | 97 ± 1 |
| Pyrene | 95 ± 1 |

**Table I 5:** Calibration curve information of the HPACs and Halo-HPACs standards (pg/µL, n=3)

| Standard name | Selected ISTD | Selected range in pg/µL | linear $R^2$ | Linear bias max (%)[a] | The Average RF | RSD (%) for RF model | Average $r_f$ range |
|---|---|---|---|---|---|---|---|
| Quinoline | Naphthalene-d8 | 2.5 - 800 | 0.9964 | +200 | 0.0058 ± 0.0006 | 10.90 | 71 -128 |
| Isoquinoline | Naphthalene-d8 | 5.0 - 800 | 0.9967 | +300 | 0.0038 ± 0.0010 | 25.73 | 55 - 122 |
| Indole | Naphthalene-d8 | 2.5 - 800 | 0.9951 | +260 | 0.0084 ± 0.0013 | 15.64 | 68-137 |
| 2-Methylquinoline | Naphthalene-d8 | 2.5 - 800 | 0.9936 | +300 | 0.0024 ± 0.0005 | 22.44 | 60 -156 |
| 4-Methylbenzothiophene | Naphthalene-d8 | 2.5 - 800 | 0.9928 | + 75 | 0.0019 ± 0.0003 | 15.87 | 64 -150 |
| 3-Methylisoquinoline | Naphthalene-d8 | 5.0 - 800 | 0.9862 | +350 | 0.0046 ± 0.0010 | 21.15 | 64-145 |
| 2,5-Dimethyl-1-phenylpyrrole | Naphthalene-d8 | 2.5-800 | 0.9898 | +200 | 0.0025 ± 0.0005 | 18.55 | 56-130 |
| 2,3-Dimethylbenzothiophene | Fluorene-d10 | 2.5 - 800 | 0.9956 | -170 | 0.0021 ± 0.0003 | 16.17 | 64-141 |
| 4-Phenylpiridine | Naphthalene-d8 | 25 - 800 | 0.9931 | +150 | 0.0031 ± 0.0006 | 19.05 | 67 -128 |
| 2,3-Dimethylindole | Naphthalene-d8 | 2.5 - 800 | 0.9905 | +310 | 0.0036 ± 0.0004 | 12.12 | 78 -139 |
| Dibenzofuran | Naphthalene-d8 | 2.5 - 800 | 0.9933 | +49 | 0.0117 ± 0.0017 | 14.27 | 72 - 139 |
| 2,3,4-Trimethylbenzothiophene | Phenanthrene-d10 | 2.5 - 800 | 0.9957 | -349 | 0.0088 ± 0.0008 | 9.24 | 83 - 123 |
| 2-Methyldibenzofuran | Fluorene-d10 | 2.5 - 800 | 0.995 | -216 | 0.0045 ± 0.0003 | 6.23 | 87 - 110 |
| Dibenzothiophene | Phenanthrene-d10 | 2.5 - 800 | 0.9985 | -175 | 0.0068 ± 0.0006 | 9.16 | 86 - 131 |
| Benzo[h]quinoline | Phenanthrene-d10 | 2.5 - 800 | 0.9988 | -63 | 0.0091 ± 0.0010 | 10.79 | 84 -143 |
| 1,2,3,4-Tetrahydrocarbazole | Phenanthrene-d10 | 2.5 - 800 | 0.9957 | +500 | 0.0012 ± 0.0002 | 17.77 | 65 -150 |
| Acridine | Phenanthrene-d10 | 2.5 - 800 | 0.9957 | +390 | 0.0056 ± 0.0010 | 18.03 | 53-140 |
| 9-Methylcarbazole | Phenanthrene-d10 | 2.5 - 800 | 0.9982 | - 200 | 0.0166 ± 0.0016 | 9.71 | 71-119 |
| Carbazole | Phenanthrene-d10 | 2.5 - 800 | 0.9917 | +422 | 0.0046 ± 0.0007 | 15.70 | 78-123 |
| 4-Methyldibenzothiophene | Phenanthrene-d10 | 2.5 - 800 | 0.9955 | -168 | 0.0034 ± 0.0002 | 7.19 | 89-111 |
| 3-Methylbenzo[f]quinoline | Phenanthrene-d10 | 2.5 - 800 | 0.9981 | +299 | 0.0021 ± 0.0004 | 21.01 | 61-120 |
| 2-Methylacridine | Phenanthrene-d10 | 2.5 - 800 | 0.9983 | +251 | 0.0041 ± 0.0010 | 24.39 | 58-120 |
| 2,4-Dimethyldibenzothiophene | Fluoranthene-d10 | 2.5 - 800 | 0.9985 | -48 | 0.0030 ± 0.0002 | 5.90 | 91 - 107 |
| 3,6-Dimethylcarbazole | Fluoranthene-d10 | 2.5 - 800 | 0.9934 | +23 | 0.0018 ± 0.0003 | 13.80 | 67 - 124 |
| 2,4,7-Trimethyldibenzothiophene | Pyrene-d10 | 2.5 - 800 | 0.9983 | no Bias | 0.0059 ± 0.0004 | 6.34 | 90-112 |
| Benzo[b]naphtho[2,3-d]furan | Pyrene-d10 | 2.5 - 800 | 0.9805 | + 498 | 0.0028 ± 0.0005 | 17.07 | 72-122 |
| 3-Tert-butyl-9H-carbazole | Pyrene-d10 | 2.5 - 800 | 0.9977 | -150 | 0.0043 ± 0.0002 | 5.76 | 88-108 |
| 9-Phenylcarbazole | Chrysene-d12 | 2.5 - 800 | 0.9925 | +245 | 0.0060 ± 0.0008 | 13.86 | 76-143 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | Pyrene-d10 | 2.5 - 800 | 0.9991 | +46 | 0.0020 ± 0.0001 | 6.30 | 83-117 |
| Benz[c]acridine | Chrysene-d12 | 2.5 - 800 | 0.9989 | +200 | 0.0060 ± 0.0006 | 10.18 | 79-133 |
| 11H-Benzo[a]carbazole | Chrysene-d12 | 2.5 - 800 | 0.9956 | -52 | 0.0008 ± 0.0001 | 17.29 | 72-140 |
| 7H-benzo[c]carbazole | Chrysene-d12 | 5.0 - 800 | 0.9967 | +50 | 0.0032 ± 0.0003 | 8.71 | 83-112 |
| 6-Methylbenzo[b]naphtho[2,1-d]thiophene | Benzo[b]fluranthene-d12 | 2.5 - 800 | 0.9905 | +138 | 0.0031 ± 0.0005 | 15.79 | 71-134 |
| 2-Phenyldibenzothiophene | Benzo[b]fluranthene-d12 | 25.0 - 800 | 0.9947 | no Bias | 0.0007 ± 0.0000 | 7.11 | 88-113 |
| Dinaphtho[1,2-b;1',2'-d]furan | Benzo[b]fluranthene-d12 | 5.0 - 800 | 0.9957 | +69 | 0.0081 ± 0.0009 | 10.56 | 82 -136 |
| Dibenz[c,h]acridine | Benzo(ghi)perylene-d14 | 5.0 - 800 | 0.9984 | no Bias | 0.0031 ± 0.0005 | 16.90 | 85-111 |
| 13(H)-Dibenzo[a,i]carbazole | Benzo(ghi)perylene-d14 | 25.0 - 800 | 0.983 | +50 | 0.0017 ± 0.0004 | 22.70 | 73 -166 |
| 2,3-Dibromothiophene | Fluorene-d10 | 2.5 - 800 | 0.9793 | -20 | 0.0007 ± 0.0001 | 13.95 | 68-126 |
| 5-Chlorobenzothiophene | Fluorene-d10 | 2.5 - 800 | 0.9987 | -100 | 0.0040 ± 0.0003 | 8.09 | 85 -118 |
| 2-Chloroquinoline | Fluorene-d10 | 2.5 - 800 | 0.0059 | -25 | 0.0074 ± 0.0006 | 7.80 | 86-114 |
| 2-Bromobenzothiophene | Fluorene-d10 | 2.5 - 800 | 0.9848 | +29 | 0.0012 ± 0.0002 | 15.66 | 68 -126 |
| 2,3-Dichlorobenzothiophene | Fluorene-d10 | 5.0 - 800 | 0.9753 | +11 | 0.0033 ± 0.0005 | 14.93 | 69-120 |
| 2-Bromoquinoline | Naphthalene-d8 | 2.5 - 800 | 0.9956 | -25 | 0.0047 ± 0.0005 | 11.46 | 81-126 |
| 5-Chloroindole | Fluorene-d10 | 2.5 - 800 | 0.9906 | +60 | 0.0050 ± 0.0007 | 13.53 | 67-124 |
| 3-Bromoisoquinoline | Fluorene-d10 | 2.5 - 800 | 0.9845 | +28 | 0.0063 ± 0.0009 | 13.79 | 76 -129 |
| 5-Bromoindol | Naphthalene-d8 | 2.5 - 800 | 0.9959 | -150 | 0.0059 ± 0.0007 | 11.81 | 82-130 |
| 2,3-Dibromobenzothiophene | Phenanthrene-d10 | 2.5 - 800 | 0.9789 | -300 | 0.0030 ± 0.0004 | 14.19 | 71-130 |
| 9-Chloroacridine | Phenanthrene-d10 | 2.5 - 800 | 0.9889 | +120 | 0.0019 ± 0.0003 | 17.48 | 66-126 |
| 2-Bromodibenzothiophene | Phenanthrene-d10 | 2.5 - 800 | 0.976 | -500 | 0.0028 ± 0.0004 | 15.08 | 74-135 |
| 3,6-Dichloro-9h-carbazole | Pyrene-d10 | 2.5 - 800 | 0.9974 | +143 | 0.0013 ± 0.0002 | 13.50 | 64-139 |
| 2,8-Dibromodibenzothiophene | Pyrene-d10 | 2.5 - 800 | 0.9844 | -73 | 0.0011 ± 0.0002 | 13.80 | 79-125 |
| 3,6-Dibromocarbazole | Chrysene-d12 | 25 - 800 | 0.9801 | +160 | 0.0006 ± 0.0000 | 14.58 | 75-124 |

[a]For the linear bias, the maximum bias (expressed as a %) obtained from the graph of expected concentration vs residual percentage is given.

**Table I 6:** Mean concentrations and percent relative standard deviation (% RSD) of analytes in the six procedural blank samples used in the method validation.

| Standard name | Concentration | % RSD |
|---|---|---|
| Quinoline | 1.8 ± 0.8 | 45 |
| Isoquinoline | 2.2 ± 0.5 | 24 |
| Indole | 0.6 ± 0.3 | 47 |
| 2-Methylquinoline | 0.5 ± 0.3 | 56 |
| 4-Methylbenzothiophene | 0.0 ± 0.0 | 56 |
| 3-Methylisoquinoline | 0.4 ± 0.2 | 46 |
| 2,5-Dimethyl-1-phenylpyrrole | 0.2 ± 0.2 | 99 |
| 2,3-Dimethylbenzothiophene | 0.5 ± 0.3 | 54 |
| 4-Phenylpiridine | 2.2 ± 0.9 | 39 |
| Dibenzofuran | 3.2 ± 0.6 | 18 |
| 2,3,4-Trimethylbenzothiophene | 0.1 ± 0.0 | 30 |
| 2-Methyldibenzofuran | 0.4 ± 0.2 | 34 |
| Dibenzothiophene | 0.3 ± 0.1 | 32 |
| Benzo[h]quinoline | 0.3 ± 0.1 | 24 |
| 1,2,3,4-Tetrahydrocarbazole | 0.9 ± 0.4 | 50 |
| Acridine | 1.1 ± 0.6 | 53 |
| 9-Methylcarbazole | 0.1 ± 0.0 | 47 |
| Carbazole | 0.5 ± 0.1 | 26 |
| 4-Methyldibenzothiophene | 0.2 ± 0.1 | 40 |
| 3-Methylbenzo[f]quinoline | 0.2 ± 0.1 | 65 |
| 2-Methylacridine | 0.6 ± 0.2 | 36 |
| 2,4-Dimethyldibenzothiophene | 0.2 ± 0.1 | 43 |
| 3,6-Dimethylcarbazole | 0.3 ± 0.2 | 81 |
| 2,4,7-Trimethyldibenzothiophene | 0.1 ± 0.0 | 20 |
| Benzo[b]naphtho[2,3-d]furan | 0.3 ± 0.1 | 55 |
| 3-Tert-butyl-9H-carbazole | 0.3 ± 0.2 | 60 |
| 9-Phenylcarbazole | 0.0 ± 0.0 | 73 |
| 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | 0.0 ± 0.0 | 70 |
| Benz[c]acridine | 0.2 ± 0.1 | 38 |
| 11H-Benzo[a]carbazole | 0.1 ± 0.1 | 51 |
| 7H-benzo[c]carbazole | 0.2 ± 0.1 | 52 |
| 6-Methylbenzo[b]naphtho[2,1-d]thiophene | 0.0 ± 0.0 | 72 |
| 2-Phenyldibenzothiophene | 0.1 ± 0.1 | 118 |
| Dinaphtho[1,2-b;1',2'-d]furan | 0.1 ± 0.1 | 63 |

| | | |
|---|---|---|
| Dibenz[c,h]acridine | 0.1 ± 0.0 | 68 |
| 13(H)-Dibenzo[a,i]carbazole | 0.1 ± 0.1 | 110 |
| 2,3-Dibromothiophene | 0.1 ± 0.1 | 113 |
| 5-Chlorobenzothiophene | 0.1 ± 0.1 | 113 |
| 2-Chloroquinoline | 0.1 ± 0.1 | 125 |
| 2-Bromobenzothiophene | 0.0 ± 0.0 | 122 |
| 2,3-Dichlorobenzothiophene | 0.0 ± 0.0 | 106 |
| 2-Bromoquinoline | 0.2 ± 0.2 | 111 |
| 5-Chloroindole | 0.1 ± 0.3 | 185 |
| 3-Bromoisoquinoline | 0.1 ± 0.2 | 132 |
| 5-Bromoindol | 0.2 ± 0.2 | 84 |
| 2,3-Dibromobenzothiophene | 0.0 ± 0.0 | 109 |
| 9-Chloroacridine | 0.1 ± 0.1 | 106 |
| 2-Bromodibenzothiophene | 0.0 ± 0.0 | 93 |
| 3,6-Dichloro-9h-carbazole | 0.0 ± 0.0 | 114 |
| 2,8-Dibromodibenzothiophene | 0.0 ± 0.0 | 123 |
| 3,6-Dibromocarbazole | 0.2 ± 0.3 | 141 |

**Table I 7:** Mean concentrations ± standard deviations (µg/g, dry weight, n=3) of 16 PAHs and 30 APAHs in sediment sample from Hamilton Harbour.

| Standard Name | Mean Concentration in µg/ g | Standard Deviation |
|---|---|---|
| Acenapthene | 1.76 | 0.09 |
| Acenapthylene | 1.58 | 0.15 |
| Anthracene | 2.04 | 0.45 |
| Benz[a]anthracene | 3.98 | 0.23 |
| Benzo[a]pyrene | 3.80 | 0.23 |
| Benzo[b]fluoranthene | 3.90 | 0.62 |
| Benzo[g,h,i]perylene | 3.49 | 0.41 |
| Benzo[k]fluoranthene | 1.69 | 0.01 |
| Chrysene | 5.78 | 0.64 |
| Dibenzo[a,h]anthracene | 0.43 | 0.05 |
| Fluoranthene | 15.11 | 0.68 |
| Fluorene | 2.62 | 0.21 |
| Indeno[1,2,3-c,d]pyrene | 3.16 | 0.38 |
| Naphthalene | 3.39 | 0.27 |
| Phenanthrene | 17.45 | 3.13 |
| Pyrene | 10.75 | 1.41 |
| 1,7-Dimethylphenanthrene | 0.03 | 0.00 |
| 1,8-Dimethylphenanthrene | 0.01 | 0.00 |
| 1-Methylnaphthalene | 0.37 | 0.04 |
| 1-Methylphenanthrene | 0.26 | 0.04 |
| 2,6-Dimethylphenanthrene | 0.04 | 0.00 |
| 2-Methylnaphthalene | 0.70 | 0.06 |
| 2-Methylphenanthrene | 0.53 | 0.08 |
| 3,6-Dimethylphenanthrene | 0.03 | 0.00 |
| 3-Methylphenanthrene | 0.49 | 0.08 |
| 9/4-Methylphenanthrene | 0.26 | 0.05 |
| C1 Chrysene | 0.62 | 0.09 |
| C2 Chrysene | 0.58 | 0.15 |
| C2 Naphthalene | 0.47 | 0.05 |
| C2 Phenanthrene | 0.19 | 0.01 |
| C3 Chrysene | 0.02 | 0.00 |
| C3 Naphthalene | 0.21 | 0.02 |
| C3 Phenanthrene | 0.14 | 0.01 |
| C4 Chrysene | 0.00 | 0.00 |
| C4 Naphthalene (Butyl) | 0.06 | 0.01 |
| C4 Phenanthrene | 0.07 | 0.01 |

| | | |
|---|---|---|
| Retene | 0.38 | 0.02 |
| C1 Benzo[a]pyrene | 1.12 | 0.16 |
| C1 Fluorene | 0.53 | 0.07 |
| C1 Pyrene | 5.10 | 0.69 |
| C2 Benzo[a]pyrene | 0.12 | 0.03 |
| C2 Fluorene | 0.69 | 0.03 |
| C2 Pyrene | 0.35 | 0.06 |
| C3 Fluorene | 0.28 | 0.02 |
| C3 Pyrene | 0.29 | 0.02 |
| C4 Fluorene | 0.15 | 0.01 |
| C4 Pyrene | 0.10 | 0.01 |

**Table I 8:** Concentrations of analytes (ng/g) in the procedural blanks (n = 3) used in the processing of Randle Reef sediments.

| Standard Name | Blank Concentration (ng/ g) |
|---|---|
| Quinoline | 1.5 |
| Isoquinoline | 1.8 |
| 2-Methylquinoline | 0.3 |
| 3-Methylisoquinoline | 0.3 |
| Benzo[h]quinoline | 0.4 |
| 3-Methylbenzo[f]quinoline | 0.2 |
| 2-Chloroquinoline | n.d. |
| 2-Bromoquinoline | n.d. |
| 3-Bromoisoquinoline | n.d. |
| 4-Phenylpyridine | 0.2 |
| Indole | 1.7 |
| 5-Chloroindole | n.d. |
| 5-Bromoindole | n.d. |
| Acridine | 1.4 |
| 2-Methylacridine | 0.2 |
| Benz[c]acridine | 0.3 |
| Dibenz[c,h]acridine | 0.1 |
| 9-Chloroacridine | n.d. |
| 9-bromoacridine | 0.9 |
| Carbazole | 3.9 |
| 9-Methylcarbazole | n.d. |
| 3,6-Dimethylcarbazole | 0.3 |
| 3-Tert-butyl-9H-carbazole | n.d. |
| 9-Phenylcarbazole | n.d. |
| 11H-Benzo[a]carbazole | 0.8 |
| 7H-benzo[c]carbazole | 0.1 |
| 13(H)-Dibenzo[a,i]carbazole | 0.4 |
| 3,6-Dichloro-9h-carbazole | 0.8 |
| 3,6-Dibromocarbazole | 0.5 |
| 4-Methylbenzothiophene | n.d. |
| 2,3-Dimethylbenzothiophene | 0.8 |
| 2,3,4-Trimethylbenzothiophene | n.d. |
| Dibenzothiophene | 1.1 |
| 4-Methyldibenzothiophene | 0.3 |
| 2,4-Dimethyldibenzothiophene | 0.1 |

| | |
|---|---|
| 2,4,7-Trimethyldibenzothiophene | 0.1 |
| 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | n.d. |
| 6-Methylbenzo[b]naphtho[2,1-d]thiophene | 0.2 |
| 2,3-Dibromothiophene | n.d. |
| 2-Bromobenzothiophene | n.d. |
| 2,3-Dichlorobenzothiophene | n.d. |
| 2,3-Dibromobenzothiophene | n.d. |
| 2-Bromodibenzothiophene | n.d. |
| 2,8-Dibromodibenzothiophene | n.d. |
| Dibenzofuran | 5.8 |
| 2-Methyldibenzofuran | 1.0 |
| Benzo[b]naphtho[2,3-d]furan | 0.6 |
| Dinaphtho[1,2-b;1',2'-d]furan | 0.5 |

# APPENDIX II. ADDITIONAL METRICS FROM CHAPTER 4: HETEROCYCLIC AROMATIC COMPOUNDS IN LAKE ONTARIO SEDIMENTS: SPATIOTEMPORAL DISTRIBUTION AND NEW APPROACHES TO FINGERPRINTING

**Table II 1:** Target analytes and internal standards used for the analysis

| **Polyaromatic Nitrogen Heterocycles** | **Polyaromatic Sulfur Heterocycles** | **Polyaromatic Oxygen Heterocycles** | **Polycyclic Aromatic Hydrocarbons** | **Internal Standards** |
|---|---|---|---|---|
| Quinoline | 4-Methylbenzothiophene | Dibenzofuran | Acenapthene | Naphthalene-d8 |
| Isoquinoline | 2,3-Dimethylbenzothiophene | 2-Methyldibenzofuran | Acenapthylene | Fluorene-d10 |
| Indole | 2,3,4-Trimethylbenzothiophene | Benzo[b]naphtho[2,3-d]furan | Anthracene | Phenanthrene-d10 |
| 2-Methylquinoline | Dibenzothiophene | Dinaphtho[1,2-b;1',2'-d]furan | Benz[a]anthracene | Anthracene-d10 |
| 3-Methylisoquinoline | 4-Methyldibenzothiophene | | Benzo[a]pyrene | Fluoranthene-d10 |
| 2,5-Dimethyl-1-phenylpyrrole | 2,4-Dimethyldibenzothiophene | | Benzo[b]fluoranthene | Pyrene-d11 |
| 4-Phenylpyridine | 2,4,7-Trimethyldibenzothiophene | | Benzo[g,h,i]perylene | Chrysene-d12 |
| 2,3-Dimethylindole | 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | | Benzo[k]fluoranthene | Benz[a]anthracene-d12 |
| Benzo[h]quinoline | 6-Methylbenzo [b]naphtho[2,1-d]thiophene | | Chrysene | Benzo[b]fluranthene-d12 |
| 1,2,3,4-Tetrahydrocarbazole | 2-Bromobenzothiophene | | Dibenzo[a,h]anthracene | Benzo[k]fluranthene-d12 |
| Acridine | 2,3-Dichlorobenzothiophene | | Fluoranthene | Benzo(ghi)perylene-d14 |
| 9-Methylcarbazole | 2,3-Dibromobenzothiophene | | Fluorene | |
| Carbazole | 2-Bromodibenzothiophene | | Indeno[1,2,3-c,d]pyrene | |
| 3-Methylbenzo[f]quinoline | 2,8-Dibromodibenzothiophene | | Naphthalene | |
| 2-Methylacridine | | | Phenanthrene | |
| 3,6-Dimethylcarbazole | | | Pyrene | |
| 2,4,7-Trimethyldibenzothiophene | | | | |
| 3-Tert-butyl-9H-carbazole | | | | |
| 9-Phenylcarbazole | | | | |
| Benz[c]acridine | | | | |
| 11H-Benzo[a]carbazole | | | | |
| Dibenz[c,h]acridine | | | | |
| 13H-Dibenzo[a,i]carbazole | | | | |
| 2-Chloroquinoline | | | | |

| | | | | |
|---|---|---|---|---|
| 2-Bromoquinoline | | | | |
| 5-Chloroindole | | | | |
| 3-Bromoisoquinoline | | | | |
| 5-Bromoindole | | | | |
| 9-Chloroacridine | | | | |
| 9-Bromoacridine | | | | |
| 3,6-Dichloro-9h-carbazole | | | | |
| 3,6-Dibromocarbazole | | | | |

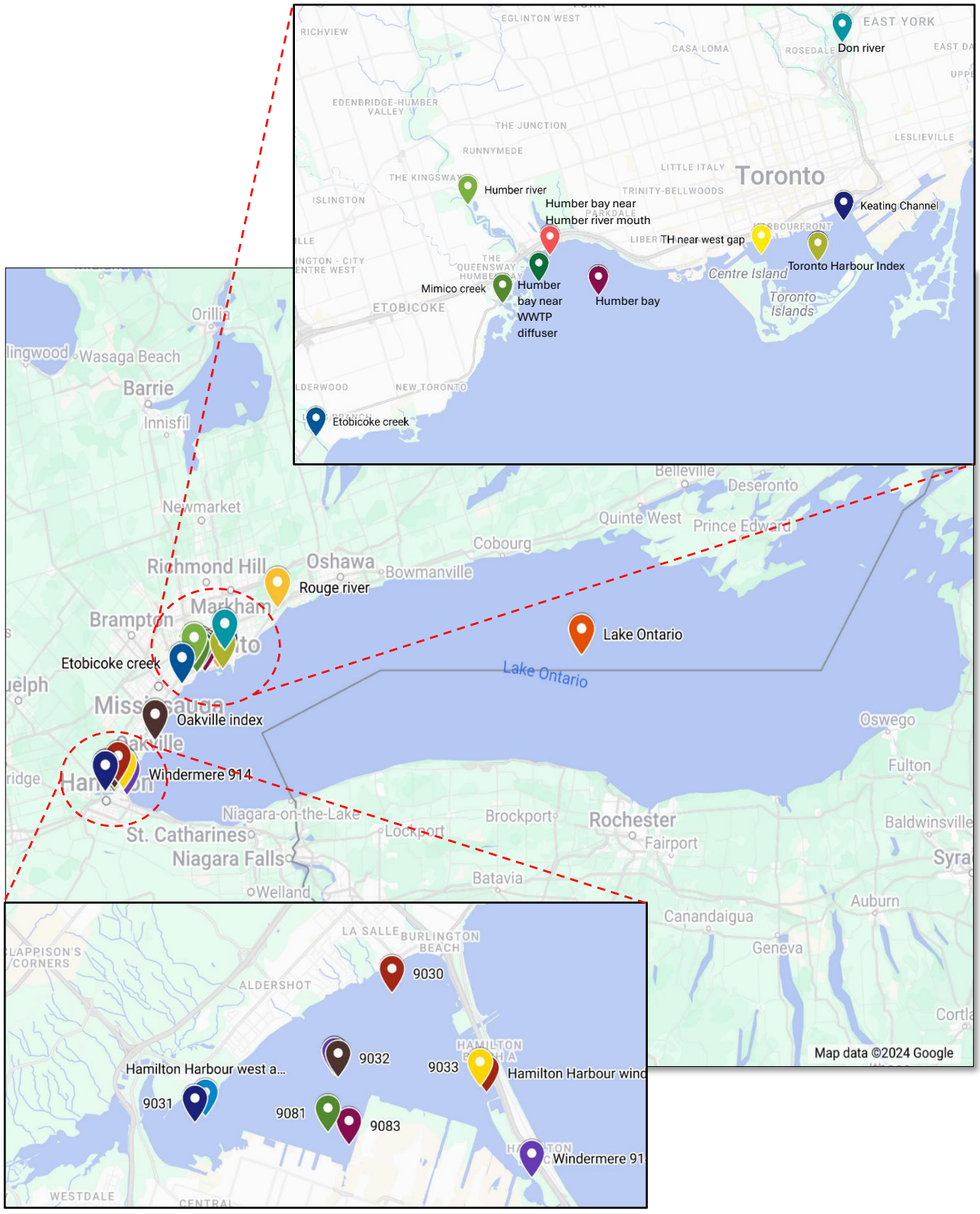



**Figure II 1**: Geographical locations of the sampling sites created using Google My Maps

**Table II 2:** Sample names and sampling information for sediments used in this study (HH - Hamilton Harbour, TH - Toronto Harbour and NH – Non Harbour)

| **Sample Name** | **Sampling Year/ Years** | **Locations** | | **Sediment Type** | **Sampling Location Name** |
|---|---|---|---|---|---|
| | | **Latitude (° N)** | **Longitude (° W)** | | |
| HH1_Year | 2014, 2017, 2018, 2019, 2022 and 2023 | 43.305 | -79.820 | Suspended Sediments | Hamilton Harbour Station 9030 |
| HH2_Year | 2014, 2017, 2018, 2019, 2022 and 2023 | 43.279 | -79.875 | Suspended Sediments | Hamilton Harbour Station 9031 |
| HH3_Year_ Depth | 2014, 2017, 2018, 2019, 2022 and 2023 | 43.288 | -79.835 | Suspended Sediments | Hamilton Harbour Station 9032 |
| HH4_Year_ Depth | 2014, 2017, 2018, 2019, 2022 and 2023 | 43.287 | -79.796 | Suspended Sediments | Hamilton Harbour Station 9033 |
| HH5_Year | 2014, 2017, 2018, 2019 and 2022 | 43.278 | -79.838 | Suspended Sediments | Hamilton Harbour Station 9081 |
| HH6_Year | 2014, 2017, 2018, 2019 and 2022 | 43.275 | -79.832 | Suspended Sediments | Hamilton Harbour Station 9083 |
| HH7_Year | 2014, 2017 and 2018 | 43.269 | -79.782 | Suspended Sediments | Windermere 914 |
| HH8_Year | 2014 and 2018 | 43.289 | -79.836 | Bottom Sediments | Hamilton Harbour Station 9032 |
| HH9_2014 | 2014 | 43.281 | -79.872 | Bottom Sediments | Hamilton Harbour Station 9031 |
| HH10_2014 | 2014 | 43.285 | -79.794 | Bottom Sediments | Hamilton Harbour Station 9033 |
| TH1_Year | 2014 and 2018 | 43.623 | -79.447 | Bottom Sediments | Toronto Harbour Index |

| | | | | | |
|---|---|---|---|---|---|
| TH2_2014 | 2014 | 43.642 | -79.361 | Bottom Sediments | Toronto Harbour near Don River/ Keating Channel |
| TH3_2014 | 2014 | 43.633 | -79.390 | Bottom Sediments | Toronto Harbour near West Gap/ Island Airport |
| NH1_2018 | 2018 | 43.805 | -79.134 | Bottom Sediments | Rouge River |
| NH2_2018 | 2018 | 43.588 | -79.546 | Bottom Sediments | Etobicoke Creek |
| NH3_2018 | 2018 | 43.686 | -79.361 | Bottom Sediments | Don River |
| NH4_2018 | 2018 | 43.426 | -79.661 | Bottom Sediments | Oakville Index |
| NH5_2018 | 2018 | 43.646 | -79.493 | Bottom Sediments | Humber River |
| NH6_2018 | 2018 | 43.621 | -79.480 | Bottom Sediments | Mimico Creek |
| NH7_Year | 2014 and 2018 | 43.623 | -79.447 | Bottom Sediments | Humber Bay Index |
| NH8_Year | 2014 and 2018 | 43.633 | -79.464 | Bottom Sediments | Humber Bay near Humber River mouth |
| NH9_2014 | 2014 | 43.627 | -79.468 | Bottom Sediments | Humber Bay near WWTP diffuser |

**Table II 3:** Detection percentages calculated from considering all the samples and reference materials

| Analyte Name | Detection Percentages % |
|---|---|
| Dibenzofuran | 98 |
| 2-Methyldibenzofuran | 99 |
| Dibenzothiophene | 100 |
| 9-Methylcarbazole | 43 |
| 4-Methyldibenzothiophene | 100 |
| 3-Methylbenzo[f]quinoline | 93 |
| 2-Methylacridine | 100 |
| 2,4-Dimethyldibenzothiophene | 95 |
| 3,6-Dimethylcarbazole | 75 |
| 2,4,7-Trimethyldibenzothiophene | 99 |
| Benzo[b]naphtho[2,3-d]furan | 100 |
| 3-Tert-butyl-9H-carbazole | 16 |
| 9-Phenylcarbazole | 6 |
| 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | 27 |
| Benz[c]acridine | 99 |
| 11H-Benzo[a]carbazole | 98 |
| 6-Methylbenzo [b]naphtho[2,1-d]thiophene | 100 |
| Dinaphtho[1,2-b;1',2'-d]furan | 98 |
| Dibenz[c,h]acridine | 94 |
| 13H-Dibenzo[a,i]carbazole | 94 |
| 9-Chloroacridine | 20 |
| 2-Bromodibenzothiophene | 0 |
| 3,6-Dichloro-9h-carbazole | 84 |
| 2,8-Dibromodibenzothiophene | 53 |
| 3,6-Dibromocarbazole | 4 |

**Table II 4:** Concentrations (ng/g, dry weight) of HPAC in sediments in each year from each sampling site

| **Sample Details** | **Year** | **$\sum_{31}$PANH** | **$\sum_{30}$PANH (without Indole)** | **$\sum_{15}$PASH** | **$\sum_{4}$PAOH** | **$\sum_{49}$HPACs (without Indole)** |
|---|---|---|---|---|---|---|
| Suspended Sediments of Hamilton Harbour Station 9030 (HH1_Year) | 2014 | 2077.5 | 439.6 | 124.7 | 259.2 | 823.4 |
| | 2017 | 3254.0 | 430.0 | 138.1 | 284.3 | 852.4 |
| | 2018 | 2956.2 | 308.8 | 87.5 | 202.4 | 598.7 |
| | 2019 | 1826.7 | 265.3 | 93.7 | 180.6 | 539.6 |
| | 2022 | 2320.0 | 354.7 | 89.6 | 234.3 | 678.7 |
| | 2023 | 1627.0 | 477.8 | 139.7 | 349.2 | 966.7 |
| Suspended Sediments of Hamilton Harbour Station 9031 (HH2_Year) | 2014 | 1295.1 | 267.4 | 75.8 | 143.7 | 486.9 |
| | 2017 | 1228.1 | 285.3 | 76.1 | 143.9 | 505.4 |
| | 2018 | 2608.4 | 268.2 | 86.6 | 163.2 | 518.0 |
| | 2019 | 2311.5 | 216.0 | 54.1 | 116.3 | 386.5 |
| | 2022 | 1865.8 | 295.2 | 94.3 | 187.0 | 576.5 |
| | 2023 | 1312.2 | 279.5 | 64.2 | 141.9 | 485.6 |
| Suspended Sediments of Hamilton Harbour Station 9032 (HH3_Year_11 m) | 2014 | 829.1 | 485.7 | 200.0 | 306.5 | 992.2 |
| | 2017 | 2246.1 | 389.6 | 119.1 | 291.5 | 800.2 |
| | 2018 | 3403.5 | 258.7 | 93.1 | 186.1 | 537.9 |
| | 2019 | 7032.5 | 235.5 | 76.3 | 124.6 | 436.4 |
| | 2022 | 3506.7 | 405.6 | 120.8 | 302.0 | 828.4 |
| | 2023 | 1849.1 | 482.6 | 143.0 | 344.5 | 970.1 |
| Suspended Sediments of Hamilton Harbour Station 9033 (HH4_Year_11 m) | 2014 | 4204.3 | 848.0 | 305.4 | 663.9 | 1817.3 |
| | 2017 | 2652.8 | 636.9 | 225.6 | 519.2 | 1381.7 |
| | 2018 | 1168.0 | 429.1 | 149.3 | 335.4 | 913.7 |
| | 2019 | 1168.0 | 429.1 | 149.3 | 335.4 | 913.7 |
| | 2022 | 2683.4 | 537.4 | 191.1 | 486.7 | 1215.2 |
| | 2023 | 2683.4 | 537.4 | 191.1 | 486.7 | 1215.2 |
| Suspended Sediments of Hamilton Harbour Station 9081 (HH5_Year) | 2014 | 600.8 | 459.8 | 162.3 | 292.5 | 914.6 |
| | 2017 | 1805.0 | 481.4 | 146.1 | 335.7 | 963.1 |
| | 2018 | 2514.3 | 271.7 | 74.6 | 191.0 | 537.3 |
| | 2019 | 2705.3 | 237.4 | 79.5 | 167.3 | 484.2 |
| | 2022 | 1919.2 | 435.3 | 119.5 | 274.3 | 829.1 |
| Suspended Sediments of Hamilton Harbour Station 9083 (HH6_Year) | 2014 | 1451.0 | 664.0 | 187.0 | 473.5 | 1324.5 |
| | 2017 | 2645.7 | 713.3 | 210.2 | 535.7 | 1459.2 |
| | 2018 | 5561.4 | 355.8 | 103.0 | 270.4 | 729.2 |
| | 2019 | 4175.8 | 347.5 | 95.9 | 230.4 | 673.7 |
| | 2022 | 2969.1 | 738.6 | 197.2 | 542.6 | 1478.3 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Suspended Sediments of Windermere 914 (HH7_Year) | 2014 | 6668.4 | 1170.6 | 215.3 | 354.2 | 1740.1 |
| | 2017 | 6071.4 | 824.8 | 161.0 | 218.3 | 1204.1 |
| | 2018 | 5643.3 | 777.3 | 247.0 | 302.7 | 1326.9 |
| Bottom Sediments of Hamilton Harbour station 9032 (HH8_Year) | 2014 | 766.1 | 752.0 | 254.4 | 506.6 | 1513.0 |
| | 2018 | 992.3 | 978.6 | 384.8 | 760.7 | 2124.0 |
| Bottom Sediments of Hamilton Harbour Station 9031 (HH9_Year) | 2014 | 413.8 | 401.5 | 122.1 | 229.6 | 753.2 |
| Bottom Sediments of Hamilton Harbour 9033 (HH10_Year) | 2014 | 1013.0 | 980.0 | 391.4 | 763.2 | 2134.5 |
| Bottom Sediments Toronto Harbour Index (TH1_Year) | 2014 | 371.4 | 362.9 | 125.4 | 161.7 | 650.0 |
| | 2018 | 504.1 | 486.3 | 180.5 | 187.8 | 854.5 |
| Bottom Sediments Toronto Harbour near Keating Channel (TH2_Year) | 2014 | 419.5 | 178.9 | 166.3 | 764.6 | 419.5 |
| Bottom Sediments Toronto Harbour near West Gap/ Island Airport (TH3_Year) | 2014 | 490.8 | 163.0 | 250.9 | 904.7 | 490.8 |
| Bottom Sediments Rouge River (NH1_Year) | 2018 | 44.9 | 42.3 | 9.4 | 7.1 | 58.8 |
| Bottom Sediments Etobicoke Creek (NH2_Year) | 2014 | 41.9 | 38.7 | 16.7 | 11.9 | 67.3 |
| Bottom Sediments Don River (NH3_Year) | 2018 | 45.4 | 44.5 | 14.9 | 9.4 | 68.7 |
| Bottom Sediments Oakville Index (NH4_Year) | 2018 | 316.4 | 310.6 | 132.8 | 174.2 | 617.6 |
| Bottom Sediments Humber River (NH5_Year) | 2018 | 145.2 | 128.9 | 41.6 | 33.1 | 203.6 |
| Bottom Sediments Mimico Creek (NH6_Year) | 2018 | 109.5 | 107.0 | 39.8 | 36.5 | 183.2 |
| Bottom Sediments Humber Bay Index (NH7_Year) | 2014 | 147.1 | 142.2 | 54.1 | 74.2 | 270.5 |
| | 2018 | 150.0 | 146.9 | 61.3 | 80.6 | 288.8 |
| Bottom Sediments Humber Bay near Humber River mouth (NH8_Year) | 2014 | 64.8 | 62.6 | 22.0 | 23.0 | 107.6 |
| | 2018 | 190.8 | 184.0 | 62.0 | 61.3 | 307.4 |
| Bottom Sediments Humber Bay near WWTP Diffuser (NH9_Year) | 2014 | 303.9 | 245.5 | 78.5 | 76.3 | 400.3 |

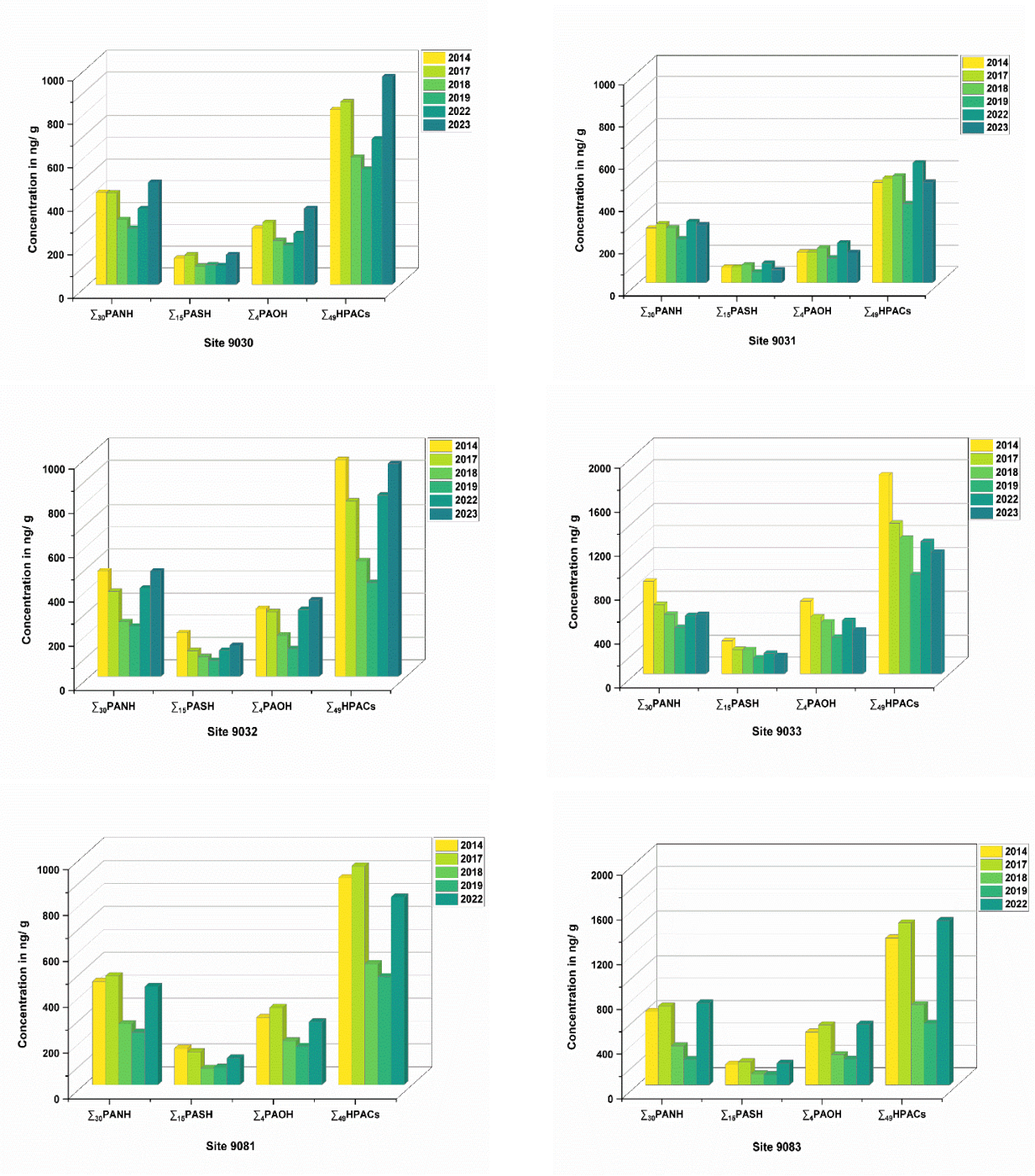



**Figure II 2:** Concentration variances of Hamilton Harbour suspended sediment samples from 2014 to 2023

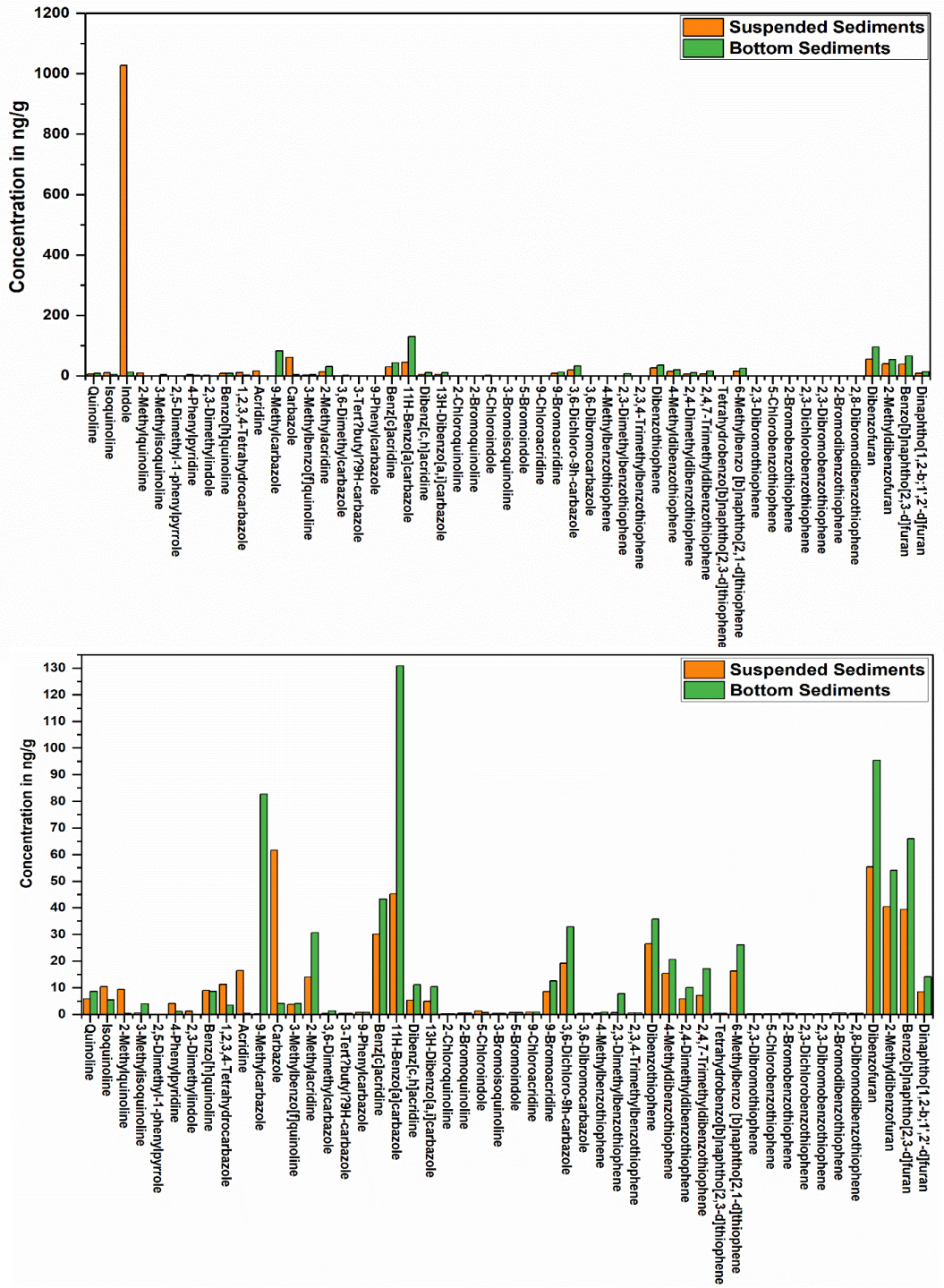



**Figure II 3:** The concentration differences of HPACs in suspended sediments from HH site 9031 (a) with indole (top panel) and (b) without indole (bottom panel)

**Table II 5:** Mass accuracies in detected analytes calculated from using five random samples

| Detected Analytes | Detected Mass | | Actual Mass | Mass Accuracy ppm |
|---|---|---|---|---|
| | Mean | SD | | |
| Quinoline | 129.0572 | 0.0001 | 129.0573 | 1 |
| Indole | 117.0573 | 0.0002 | 117.0573 | 0 |
| Dibenzofuran | 168.0569 | 0.0002 | 168.057 | 1 |
| Dibenzothiophene | 184.0341 | 0.0002 | 184.0341 | 0 |
| Carbazole | 167.0728 | 0.0002 | 167.073 | 1 |
| 4-Methyldibenzothiophene | 198.0494 | 0.0002 | 198.0498 | 2 |
| 2,4-Dimethyldibenzothiophene | 212.0651 | 0.0003 | 212.0654 | 2 |
| 2,4,7-Trimethylbenzothiophene | 226.0809 | 0.0001 | 226.0811 | 1 |
| Benzo[b]naphtho[2,3-d]furan | 218.0727 | 0.0004 | 218.0726 | 0 |
| Benz[c]acridine | 229.0884 | 0.0004 | 229.0886 | 1 |
| 11H-Benzo(c)carbazole | 217.0885 | 0.0004 | 217.0886 | 0 |
| 6-Methylbenzo[b]naphtho[2,1-d]thiophene | 248.0645 | 0.0001 | 248.0654 | 4 |
| Dinaphtho[1,2-b:1',2'-d]furan | 268.0884 | 0.0005 | 268.0883 | 0 |
| 7H-Benzo[c]carbazole | 217.0886 | 0.0004 | 217.0886 | 0 |

**Table II 6:** Summary of screening criteria for analytes suitable for chemical fingerprinting

| Analytes | log Kow | Water solubility (mg/L) | Sorption constant log Koc | Half live in sediments (hr)* | 9031 Suspended sediments (ng/g) | 9031 bottom sediments (ng/g) | Concentration criteria (Yes/ No) | Slope analysis criteria (Yes/ No) | # Rings | Removed/ Retained for Fingerprinting |
|---|---|---|---|---|---|---|---|---|---|---|
| Quinoline | 2.03 | 6110.000 | 3.1 | 3240 | 5.9 | 8.7 | YES | NO | 2 | Removed |
| Isoquinoline | 2.08 | 5663.000 | 3.19 | 3240 | 10.4 | 5.5 | NO | YES | 2 | Removed |
| Indole | 2.14 | 3560.000 | 2.93 | 3240 | 1027.7 | 12.3 | NO | YES | 2 | Removed |
| 2-Methylquinoline | 2.59 | 818.270 | 3.39 | 3240 | 9.3 | 0.4 | NO | YES | 2 | Removed |
| 3-Methylisoquinoline | 2.68 | 269.710 | 3.39 | 3240 | 0.6 | 4.0 | YES | YES | 2 | Removed |
| 2,5-Dimethyl-1-phenylpyrrole | 3.74 | 39.014 | 3.88 | 8100 | 0.0 | 0.0 | NO | NO | 2 | Removed |
| 4-Phenylpyridine | 2.59 | 267.400 | 3.4 | 8100 | 4.0 | 1.2 | NO | NO | 2 | Removed |
| 2,3-Dimethylindole | 3.15 | 53.191 | 3.36 | 8100 | 1.2 | 0.0 | NO | NO | 2 | Removed |
| Benzo[h]quinoline | 3.43 | 78.700 | 4.64 | 3240 | 9.1 | 8.7 | NO | NO | 3 | Removed |
| 1,2,3,4-Tetrahydrocarbazole | 4.01 | 4.534 | 3.96 | 8100 | 11.3 | 3.5 | NO | NO | 3 | Removed |
| Acridine | 3.48 | 38.400 | 4.18 | 3240 | 16.5 | 0.4 | NO | NO | 3 | Removed |
| 9-Methylcarbazole | 3.78 | 14.563 | 2.93 | 8100 | 0.2 | 82.8 | YES | YES | 3 | Retained |
| Carbazole | 3.29 | 1.800 | 2.92 | 8100 | 61.6 | 4.1 | NO | YES | 3 | Removed |
| 3-Methylbenzo[f]quinoline | 3.87 | 1.836 | 4.43 | 8100 | 3.8 | 4.1 | YES | YES | 3 | Retained |
| 2-Methylacridine | 3.87 | 1.830 | 4.42 | 8100 | 14.0 | 30.7 | YES | YES | 3 | Retained |
| 3,6-Dimethylcarbazole | 4.32 | 0.726 | 4.37 | 8100 | 0.4 | 1.4 | YES | YES | 3 | Retained |
| 3-Tert?butyl?9H-carbazole | 5.14 | 0.105 | 3.54 | 8100 | 0.4 | 0.4 | YES | YES | 3 | Retained |
| 9-Phenylcarbazole | 4.99 | 0.108 | 4.48 | 8100 | 0.8 | 0.8 | YES | YES | 4 | Retained |
| Benz[c]acridine | 4.48 | 0.350 | 5.25 | 13000 | 30.1 | 43.3 | YES | YES | 4 | Retained |
| 11H-Benzo[a]carbazole | 4.41 | 1.301 | 5 | 13000 | 45.3 | 130.9 | YES | YES | 4 | Retained |
| Dibenz[c,h]acridine | 5.67 | 4.000 | 6.28 | 13000 | 5.4 | 11.2 | YES | YES | 5 | Retained |
| 13H-Dibenzo[a,i]carbazole | 5.58 | 0.008 | 6.03 | 13000 | 5.0 | 10.5 | YES | YES | 5 | Retained |
| 2-Chloroquinoline | 2.78 | 469.760 | 3.39 | 8100 | 0.3 | 0.3 | YES | YES | 2 | Removed |
| 2-Bromoquinoline | 3.03 | 102.250 | 3.39 | 8100 | 0.5 | 0.5 | YES | YES | 2 | Removed |
| 5-Chloroindole | 2.70 | 47.827 | 3.13 | 8100 | 1.3 | 0.7 | NO | NO | 2 | Removed |
| 3-Bromoisoquinoline | 3.03 | 102.250 | 3.39 | 8100 | 0.4 | 0.4 | YES | NO | 2 | Removed |
| 5-Bromoindole | 2.94 | 77.942 | 3.13 | 8100 | 0.7 | 0.7 | YES | YES | 2 | Removed |
| 9-Chloroacridine | 3.96 | 4.614 | 4.44 | 8100 | 0.9 | 0.9 | YES | YES | 3 | Retained |
| 9-Bromoacridine | 4.83 | 1.584 | 4.44 | 8100 | 8.6 | 12.6 | YES | NO | 3 | Removed |
| 3,6-Dichloro-9h-carbazole | 4.52 | 0.302 | 4.37 | 8100 | 19.2 | 32.8 | YES | YES | 3 | Retained |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 3,6-Dibromocarbazole | 5.01 | 0.036 | 4.37 | 13000 | 0.4 | 0.4 | YES | YES | 3 | Retained |
| 4-Methylbenzothiophene | 3.54 | 39.806 | 3.14 | 3240 | 0.5 | 0.9 | YES | YES | 2 | Removed |
| 2,3-Dimethylbenzothiophene | 4.08 | 21.749 | 3.36 | 8100 | 0.7 | 7.8 | YES | YES | 2 | Removed |
| 2,3,4-Trimethylbenzothiophene | 4.63 | 6.370 | 3.57 | 8100 | 0.6 | 0.6 | YES | YES | 2 | Removed |
| Dibenzothiophene | 4.59 | 1.470 | 3.57 | 8100 | 26.5 | 35.8 | YES | YES | 3 | Retained |
| 4-Methyldibenzothiophene | 3.53 | 39.806 | 3.57 | 8100 | 15.4 | 20.7 | YES | YES | 3 | Retained |
| 2,4-Dimethyldibenzothiophene | 4.08 | 21.749 | 3.57 | 8100 | 5.9 | 10.2 | YES | YES | 3 | Retained |
| 2,4,7-Trimethyldibenzothiophene | 5.81 | 0.027 | 3.57 | 8100 | 7.2 | 17.1 | YES | YES | 3 | Retained |
| 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | 6.13 | 0.012 | 4.99 | 8100 | 0.4 | 0.4 | YES | YES | 4 | Retained |
| 6-Methylbenzo [b]naphtho[2,1-d]thiophene | 5.89 | 0.018 | - | - | 16.3 | 26.1 | YES | YES | 4 | Retained |
| 2,3-Dibromothiophene | 3.59 | 84.768 | 2.32 | 8100 | 0.3 | 0.3 | YES | YES | 1 | Removed |
| 5-Chlorobenzothiophene | 4.63 | 49.210 | 3.13 | 8100 | 0.3 | 0.3 | YES | YES | 2 | Removed |
| 2-Bromobenzothiophene | 3.88 | 18.214 | 3.13 | 8100 | 0.4 | 0.4 | YES | YES | 2 | Removed |
| 2,3-Dichlorobenzothiophene | 4.28 | 9.370 | 3.36 | 8100 | 0.3 | 0.3 | YES | YES | 2 | Removed |
| 2,3-Dibromobenzothiophene | 4.77 | 1.156 | 3.36 | 8100 | 0.3 | 0.3 | YES | YES | 2 | Removed |
| 2-Bromodibenzothiophene | 5.05 | 0.075 | 4.17 | 8100 | 0.6 | 0.6 | YES | YES | 3 | Retained |
| 2,8-Dibromodibenzothiophene | 5.95 | 0.004 | 4.37 | 13000 | 0.4 | 0.4 | YES | YES | 3 | Retained |
| Dibenzofuran | 4.12 | 3.100 | 3.96 | 3240 | 55.4 | 95.5 | YES | YES | 3 | Retained |
| 2-Methyldibenzofuran | 4.26 | 0.956 | 3.13 | 3240 | 40.5 | 54.1 | YES | YES | 3 | Retained |
| Benzo[b]naphtho[2,3-d]furan | 5.75 | 0.182 | 4.99 | 8100 | 39.4 | 66.0 | YES | YES | 4 | Retained |
| Dinaphtho[1,2-b;1',2'-d]furan | 6.89 | 0.010 | 6 | 13000 | 8.5 | 14.1 | YES | YES | 5 | Retained |

* based on level III Fugacity model (MCI method)

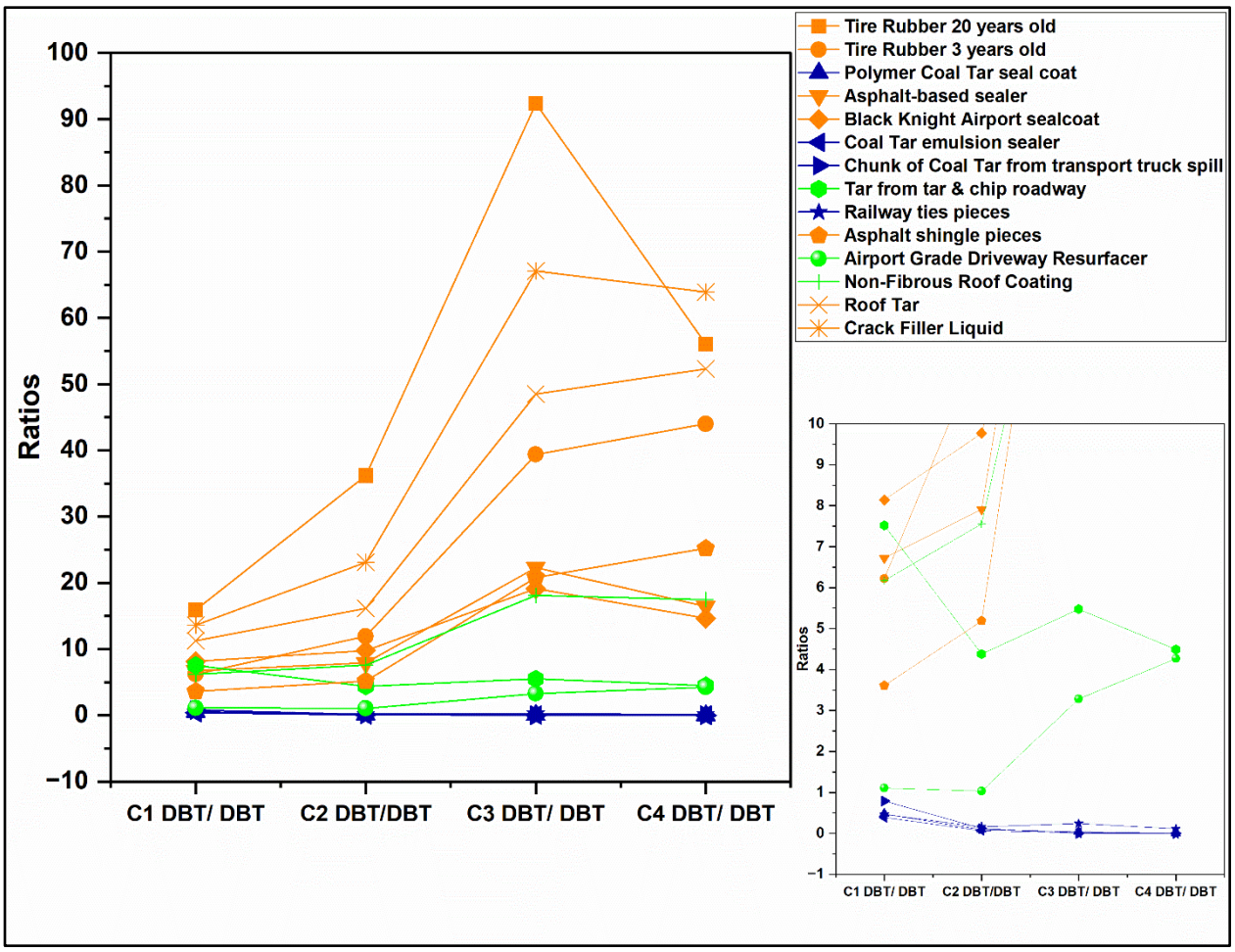



**Figure II 4:** DBT ratios calculated for the reference samples (CT RM - blue, NCT RM -orange and Mixed RM - green).

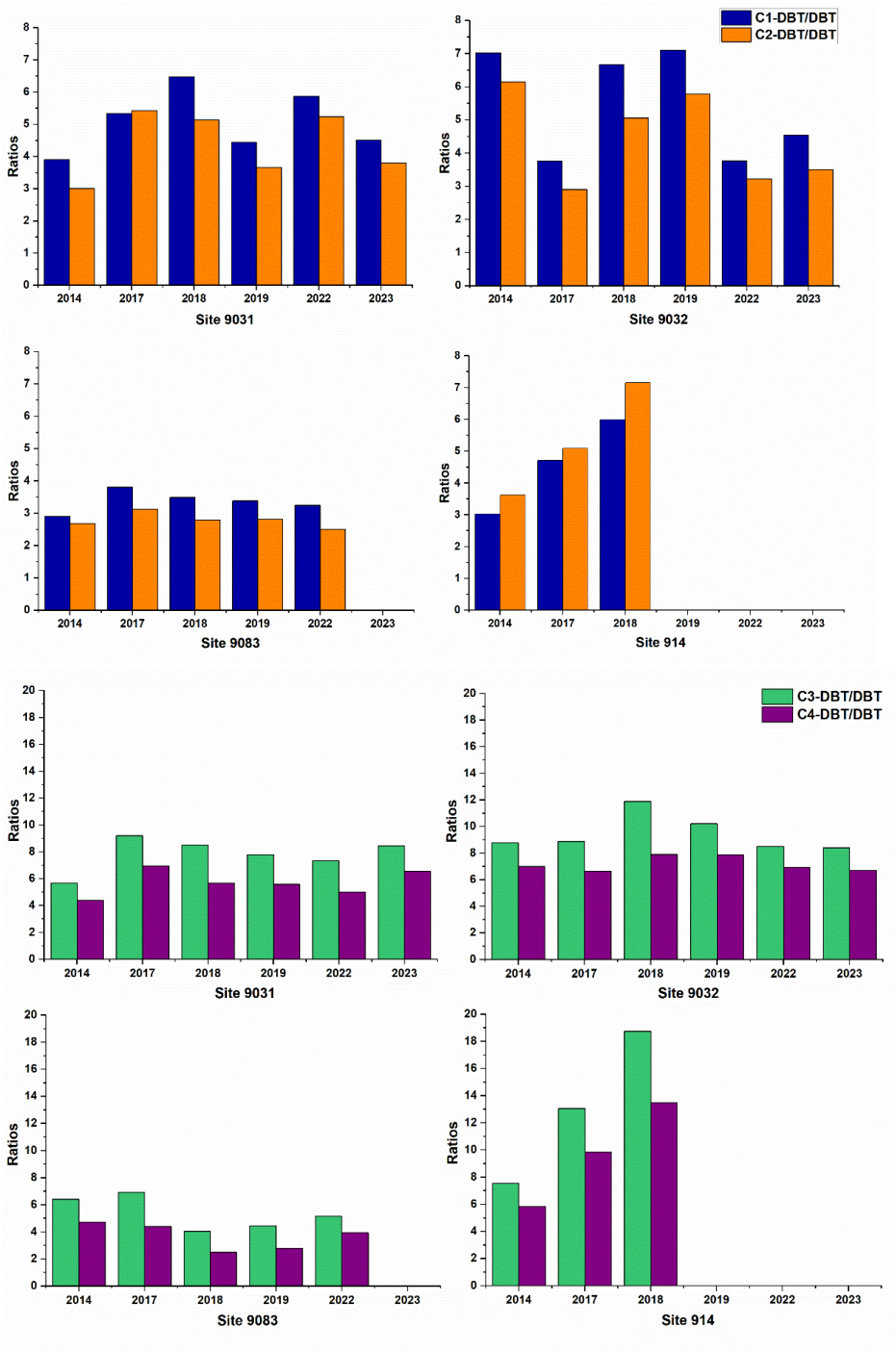



**Figure II 5:** Ratios between dibenzothiophene and its alkylated derivatives in HH suspended sediments

**Table II 7:** The Spearman's rank correlation rho values, p-values and $R^2$ values for each analyte and HPACs type

| Analyte | Spearman's rho | p-value | $R^2$ |
|---|---|---|---|
| 11H-Benzo[a]carbazole | 0.986 | 0.000 | 0.960 |
| Benz[c]acridine | 0.979 | 0.000 | 0.922 |
| Benzo[b]naphtho[2,3-d]furan | 0.978 | 0.000 | 0.956 |
| Dinaphtho[1,2-b;1',2'-d]furan | 0.962 | 0.000 | 0.915 |
| Dibenz[c,h]acridine | 0.945 | 0.000 | 0.884 |
| 13H-Dibenzo[a,i]carbazole | 0.944 | 0.000 | 0.886 |
| Dibenzothiophene | 0.941 | 0.000 | 0.865 |
| 6-Methylbenzo [b]naphtho[2,1-d]thiophene | 0.925 | 0.000 | 0.795 |
| 2-Methylacridine | 0.920 | 0.000 | 0.761 |
| 2,4-Dimethyldibenzothiophene | 0.915 | 0.000 | 0.718 |
| 2,4,7-Trimethyldibenzothiophene | 0.914 | 0.000 | 0.752 |
| 2,8-Dibromodibenzothiophene | 0.900 | 0.000 | 0.850 |
| 4-Methyldibenzothiophene | 0.842 | 0.000 | 0.553 |
| 2-Methyldibenzofuran | 0.766 | 0.000 | 0.501 |
| Dibenzofuran | 0.741 | 0.000 | 0.525 |
| 9-Bromoacridine | 0.714 | 0.000 | 0.415 |
| 3,6-Dimethylcarbazole | 0.613 | 0.000 | 0.249 |
| 3-Tert-butyl-9H-carbazole | 0.361 | 0.010 | 0.198 |
| 3,6-Dichloro-9h-carbazole | 0.337 | 0.017 | 0.077 |
| 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | 0.325 | 0.021 | 0.143 |
| 9-Methylcarbazole | 0.325 | 0.021 | 0.139 |
| 2-Bromodibenzothiophene | NA | NA | 0.500 |

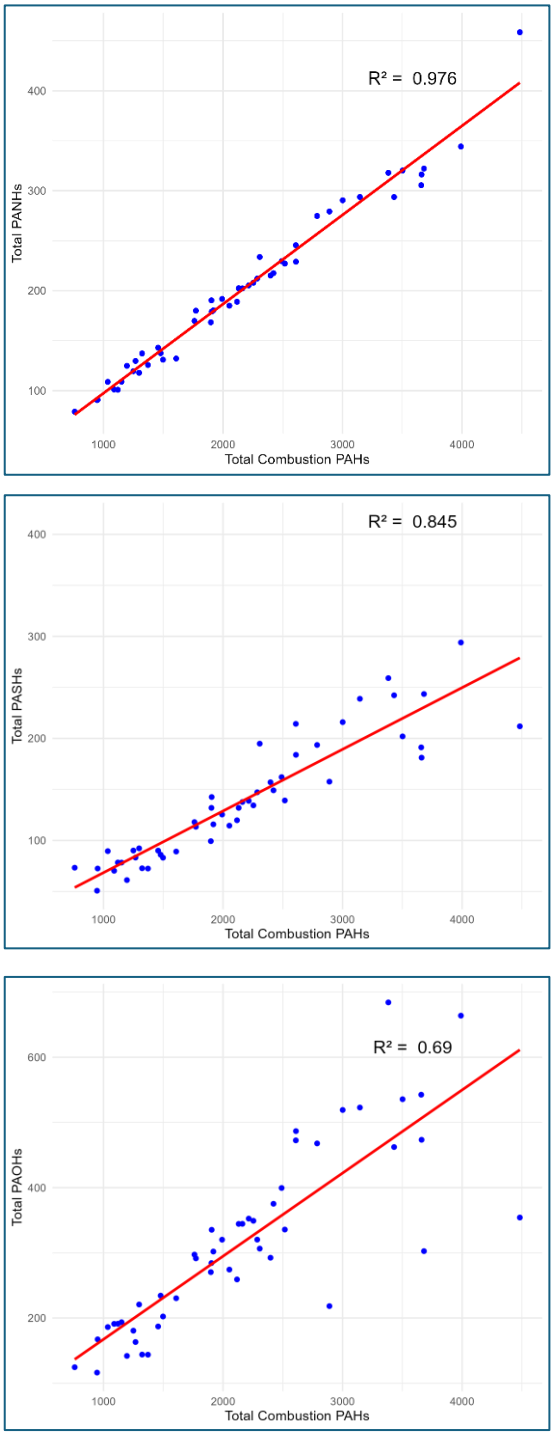



**Figure II 6:** Scatter plots between the total combustion PAHs concentration and concentrations of (a) PANHs (top panel), (b) PASHs (middle panel) and (c) PAOHs (bottom panel) in each sediment sample.

**Table II 8:** Screening criteria for the groupings of reference materials

| **Reference Material** | **Carbazole concentration ng/ g** | **$\sum_{16}$PAH concentrations ng/ g** | **C1 DBT/ DBT** | **C2 DBT/ DBT** | **C3 DBT/ DBT** | **C4 DBT/ DBT** | **Type[a]** |
|---|---|---|---|---|---|---|---|
| Tire Rubber -~20 years old | 0.55 | 84621 | 15.91 | 36.18 | 92.38 | 56.05 | NCT RM |
| Tire Rubber - ~3 years old | 0.55 | 61564 | 6.22 | 11.92 | 39.39 | 44.02 | NCT RM |
| Asphalt-based sealer1 | 0.55 | 25072 | 6.72 | 7.91 | 22.32 | 16.45 | NCT RM |
| Asphalt-based sealcoat2 | 13 | 21489 | 8.14 | 9.77 | 19.13 | 14.63 | NCT RM |
| Asphalt shingle pieces | 0.55 | 28570 | 3.61 | 5.19 | 20.81 | 25.2 | NCT RM |
| Roof Tar (used in schools) | 0.55 | 98765 | 11.26 | 16.16 | 48.51 | 52.29 | NCT RM |
| Crack Filler (liquid) | 0.55 | 22873 | 13.64 | 23.09 | 67.05 | 63.91 | NCT RM |
| Tar from tar & chip roadway (2015) - Asphalt based | 122 | 271885 | 7.52 | 4.38 | 5.48 | 4.49 | Mixed RM |
| Airport Grade Driveway Re-surfacer (Repair) | 741 | 452782 | 1.11 | 1.03 | 3.28 | 4.27 | Mixed RM |
| Non-Fibrous Roof Coating | 24 | 158157 | 6.17 | 7.56 | 18.09 | 17.47 | Mixed RM |
| Polymer Coal Tar seal coat | 1405726 | 300844702 | 0.48 | 0.09 | 0.04 | 0 | CT RM |
| Coal Tar emulsion sealer | 2146902 | 44432827 | 0.39 | 0.07 | 0.02 | 0 | CT RM |
| Chunk of Coal Tar - from transport truck spill (2015) | 1153040 | 61917534 | 0.79 | 0.12 | 0 | 0 | CT RM |
| Railway ties pieces | 1322367 | 363077747 | 0.47 | 0.16 | 0.24 | 0.11 | CT RM |

[a]NCT RM = non coal-tar based reference materials, CT RM = coal-tar based reference materials

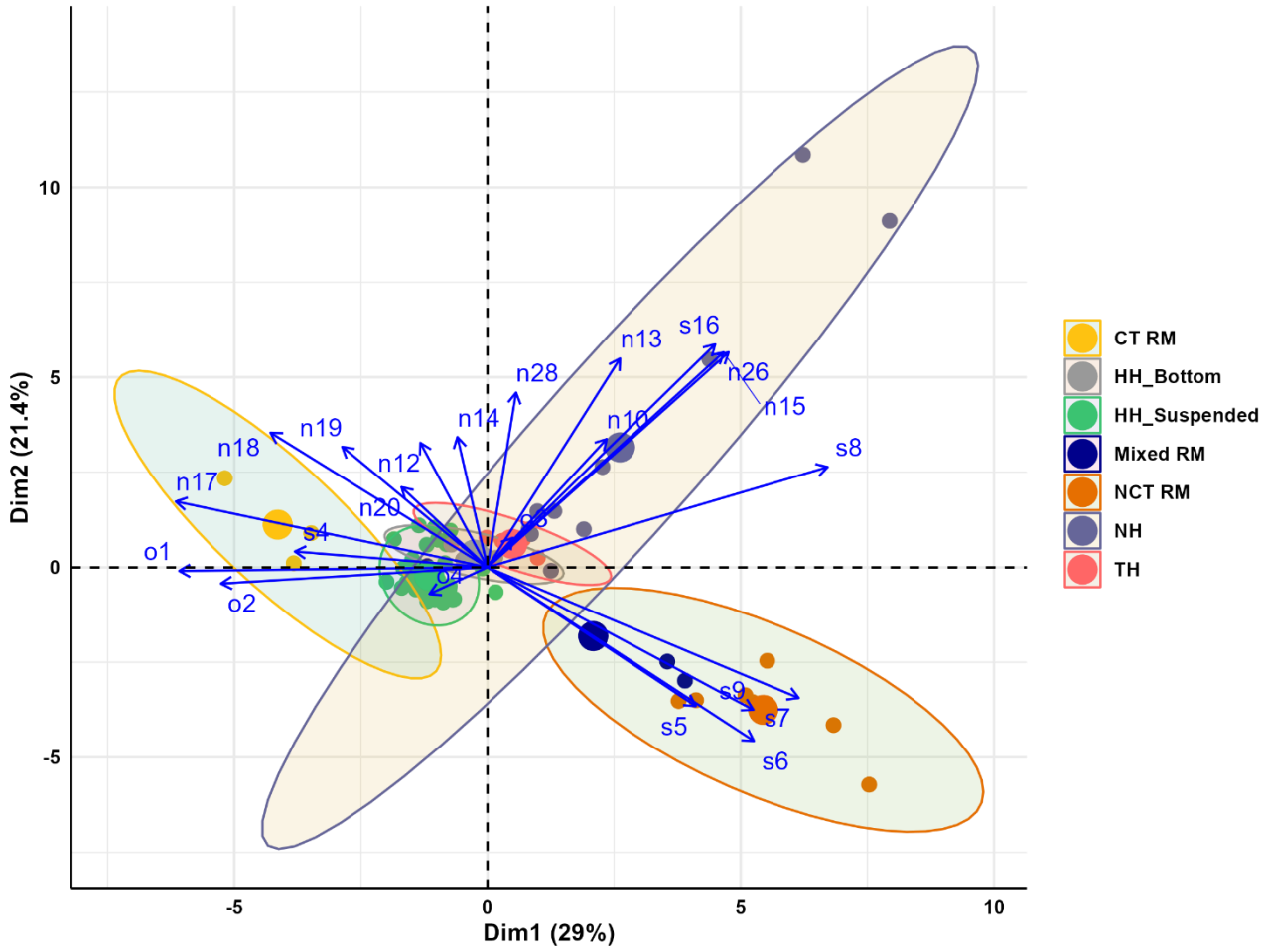



**Figure II 7:** Biplot for the principal component analysis

(Quinoline:N1, Isoquinoline:N2, Indole:N3, 2-Methylquinoline:N4, 4-Methylbenzothiophene:S1, 3-Methylisoquinoline:N5, 2,3-Dimethylbenzothiophene:S2, 4-Phenylpyridine:N6, Dibenzofuran:O1, 2,3,4-Trimethylbenzothiophene:S3, 2-Methyldibenzofuran:O2, Dibenzothiophene:S4, Benzo[h]quinoline:N7, 1,2,3,4-Tetrahydrocarbazole:N8, Acridine:N9, 9-Methylcarbazole:N10, Carbazole:N11, 4-Methyldibenzothiophen:S5, 3-Methylbenzo[f]quinoline: N12, 2-Methylacridine:N13, 2,4-Dimethyldibenzothiophene:S6, 3,6-Dimethylcarbazole:N14, 2,4,7-Trimethyldibenzothiophene:S7, Benzo[b]naphtho[2,3-d]furan:O3, 3-Tert-butyl-9H-carbazole:N15, 9-Phenylcarbazole:N16, 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene:S8, Benz[c]acridine:N17, 11H-Benzo[a]carbazole:N18, 6-Methylbenzo [b]naphtho[2,1-d]thiophene:S9, Dinaphtho[1,2-b;1',2'-d]furan:O4, Dibenz[c,h]acridine:N19, 13H-Dibenzo[a,i]carbazole:N20, 2,3-Dibromothiophene:S10, 5-Chlorobenzothiophene:S11, 2-Chloroquinoline:N21, 2-Bromobenzothiophene:S12, 2,3-Dichlorobenzothiophene:S13, 2-Bromoquinoline:N22, 5-Chloroindole:N23, 3-Bromoisoquinoline:N24, 5-Bromoindole:N25, 2,3-Dibromobenzothiophene:S14, 9-Chloroacridine:N26, 2-Bromodibenzothiophene:S15, 9-Bromoacridine:N27, 3,6-Dichloro-9h-carbazole:N28, 2,8-Dibromodibenzothiophene:S16, 3,6-Dibromocarbazole: N29)

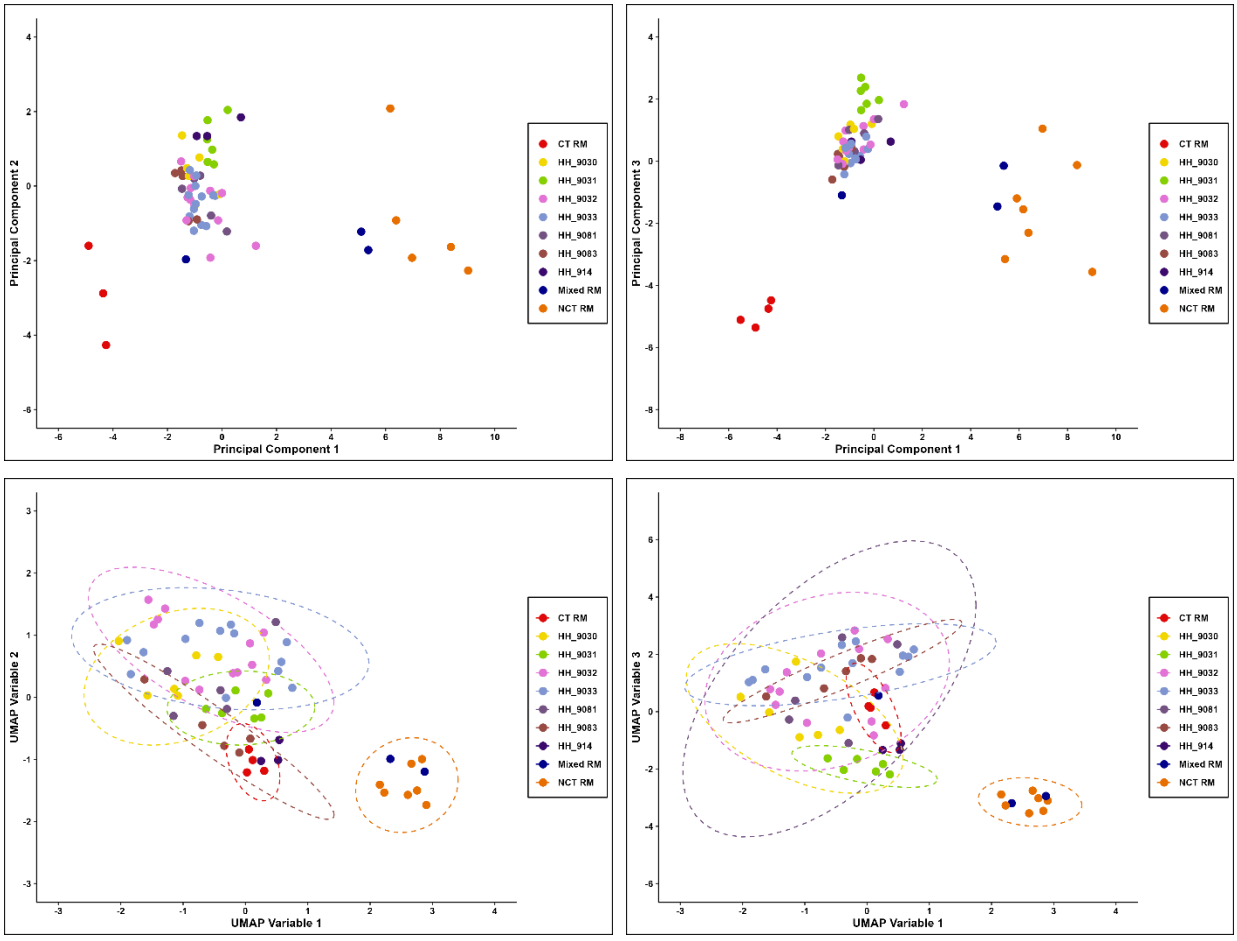



**Figure II 8:** PCA and UMAP analysis outputs between (a) PC1 and PC2 (top panel), (b) PC1 and PC3 (bottom panel) for HH harbour suspended sediments colored based on different sampling sites, the CT reference materials (yellow dots), the NCT reference materials (orange dots) and the Mixed RM (dark blue dots).

**Figure II 9:** PCA Loading values for each analyte

| Selected Analytes | PC1 | PC2 | PC3 | PC4 | PC5 |
|---|---|---|---|---|---|
| Dibenzofuran | -0.786 | -0.012 | 0.436 | -0.225 | 0.128 |
| 2-Methyldibenzofuran | -0.680 | -0.055 | 0.399 | -0.421 | 0.315 |
| Dibenzothiophene | -0.491 | 0.054 | 0.595 | 0.285 | -0.438 |
| 9-Methylcarbazole | 0.304 | 0.436 | -0.466 | -0.245 | -0.587 |
| 4-Methyldibenzothiophene | 0.531 | -0.472 | 0.386 | -0.035 | 0.174 |
| 3-Methylbenzo[f]quinoline | -0.171 | 0.423 | 0.585 | 0.267 | -0.058 |
| 2-Methylacridine | 0.338 | 0.710 | 0.356 | -0.241 | -0.142 |
| 2,4-Dimethyldibenzothiophene | 0.680 | -0.590 | 0.163 | 0.133 | 0.100 |
| 3,6-Dimethylcarbazole | -0.077 | 0.443 | 0.314 | 0.478 | 0.146 |
| 2,4,7-Trimethyldibenzothiophene | 0.794 | -0.444 | 0.132 | 0.060 | 0.113 |
| Benzo[b]naphtho[2,3-d]furan | 0.065 | 0.100 | 0.155 | -0.673 | 0.516 |
| 3-Tert-butyl-9H-carbazole | 0.602 | 0.730 | 0.149 | 0.007 | 0.124 |
| 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | 0.868 | 0.341 | 0.105 | 0.176 | 0.104 |
| Benz[c]acridine | -0.795 | 0.224 | -0.015 | 0.264 | 0.085 |
| 11H-Benzo[a]carbazole | -0.554 | 0.458 | -0.392 | 0.286 | -0.262 |
| 6-Methylbenzo [b]naphtho[2,1-d]thiophene | 0.678 | -0.485 | -0.227 | 0.332 | 0.082 |
| Dinaphtho[1,2-b;1',2'-d]furan | -0.148 | -0.091 | -0.735 | -0.084 | 0.181 |
| Dibenz[c,h]acridine | -0.371 | 0.409 | -0.202 | 0.214 | 0.584 |
| 13H-Dibenzo[a,i]carbazole | -0.220 | 0.273 | -0.388 | 0.451 | 0.588 |
| 9-Chloroacridine | 0.615 | 0.731 | 0.141 | 0.022 | 0.116 |
| 3,6-Dichloro-9h-carbazole | 0.073 | 0.594 | -0.371 | -0.468 | -0.031 |
| 2,8-Dibromodibenzothiophene | 0.581 | 0.758 | 0.064 | 0.023 | 0.084 |

# APPENDIX III. ADDITIONAL METRICS FROM CHAPTER 5: USING CHEMOMETRICS TO DISCERN PATTERNS AND SPATIAL DISTRIBUTION OF HETEROCYCLIC AROMATIC COMPOUNDS IN LAKE ERIE FISH

**Table III 1:** The targeted analytes and their reference codes used in this study

| **Polyaromatic Nitrogen Heterocycles** | **Polyaromatic Sulfur Heterocycles** | **Polyaromatic Oxygen Heterocycles** | **Polycyclic Aromatic Hydrocarbons** | **Alkylated Polycyclic Aromatic Hydrocarbons** | **Internal Standards** |
|---|---|---|---|---|---|
| Quinoline - PANH1 | 4-Methylbenzothiophene - PASH1 | Dibenzofuran - PAOH1 | Naphthalene- PAH1 | 1-Methylnaphthalene - APAH1 | Naphthalene-d8 |
| Isoquinoline - PANH2 | 2,3-Dimethylbenzothiophene - PASH2 | 2-Methyldibenzofuran - PAOH2 | Acenapthene- PAH2 | 2-Methylnaphthalene - APAH2 | Fluorene-d10 |
| Indole - PANH3 | 2,3,4-Trimethylbenzothiophene - PASH3 | Benzo[b]naphtho[2,3-d]furan - PAOH3 | Acenapthylene- PAH3 | C2 Naphthalene - APAH3 | Phenanthrene-d10 |
| 2-Methylquinoline - PANH4 | 4-Methyldibenzothiophene - PASH4 | Dinaphtho[1,2-b;1',2'-d]furan - PAOH4 | Fluorene- PAH4 | C3-Naphthalene - APAH4 | Anthracene-d10 |
| 3-Methylisoquinoline - PANH5 | 2,4-Dimethyldibenzothiophene - PASH5 | | Phenanthrene- PAH5 | C4-Naphthalene - APAH5 | Fluoranthene-d10 |
| 2,5-Dimethyl-1-phenylpyrrole - PANH6 | 2,4,7-Trimethyldibenzothiophene - PASH6 | | Anthracene- PAH6 | C1 Fluorene- APAH6 | Pyrene-d11 |
| 4-Phenylpyridine - PANH7 | 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene - PASH7 | | Fluoranthene- PAH7 | C2 Fluorene- APAH7 | Chrysene-d12 |
| 2,3-Dimethylindole - PANH8 | 6-Methylbenzo [b]naphtho[2,1-d]thiophene - PASH8 | | Pyrene- PAH8 | C3 Fluorene- APAH8 | Benz[a]anthracene-d12 |
| Benzo[h]quinoline - PANH9 | Dibenzothiophene- PASH9 | | Benz[a]anthracene- PAH9 | C4 Fluorene- APAH9 | Benzo[b]fluranthene-d12 |
| 1,2,3,4-Tetrahydrocarbazole - PANH10 | C1 Dibenzothiophene- PASH10 | | Chrysene- PAH10 | 1,7-Dimethylphenanthrene - APAH10 | Benzo[k]fluranthene-d12 |
| Acridine - PANH11 | C2 Dibenzothiophene- PASH11 | | Benzo[a]pyrene- PAH11 | 1,8-Dimethylphenanthrene - APAH11 | Benzo(ghi)perylene-d14 |
| 9-Methylcarbazole - PANH12 | C3 Dibenzothiophene- PASH12 | | Benzo[b]fluoranthene- PAH12 | 1-Methylphenanthrene - APAH12 | |
| Carbazole - PANH13 | C4 Dibenzothiophene- PASH13 | | Benzo[k]fluoranthene- PAH13 | 2,6-Dimethylphenanthrene - APAH13 | |
| 3-Methylbenzo[f]quinoline - PANH14 | 2,3-Dibromothiophene - PASH14 | | Indeno[1,2,3-c,d]pyrene- PAH14 | 2-Methylphenanthrene - APAH14 | |
| 2-Methylacridine - PANH15 | 5-Chlorobenzothiophene - PASH15 | | Dibenzo[a,h]anthracene- PAH15 | 3,6-Dimethylphenanthrene - APAH15 | |
| 3,6-Dimethylcarbazole - PANH16 | 2-Bromobenzothiophene - PASH16 | | Benzo[g,h,i]perylene- PAH16 | 3-Methylphenanthrene - APAH16 | |

| | | | | | |
|---|---|---|---|---|---|
| 3-Tert-butyl-9H-carbazole - PANH17 | 2,3-Dichlorobenzothiophene - PASH17 | | | 9/4-Methylphenanthrene - APAH17 | |
| 9-Phenylcarbazole - PANH18 | 2,3-Dibromobenzothiophene - PASH18 | | | C2 Phenanthrene - APAH18 | |
| Benz[c]acridine - PANH19 | 2-Bromodibenzothiophene - PASH19 | | | C3-Phenanthrene - APAH19 | |
| 11H-Benzo[a]carbazole - PANH20 | 2,8-Dibromodibenzohiophene - PASH20 | | | C4- Phenanthrene - APAH20 | |
| Dibenz[c,h]acridine - PANH21 | | | | C1 Pyrene- APAH21 | |
| 13H-Dibenzo[a,i]carbazole - PANH22 | | | | C2 Pyrene- APAH22 | |
| 2-Chloroquinoline - PANH23 | | | | C3 Pyrene- APAH23 | |
| 2-Bromoquinoline - PANH24 | | | | C4 Pyrene- APAH24 | |
| 5-Chloroindole - PANH25 | | | | C1 Benzo[a]pyrene- APAH25 | |
| 3-Bromoisoquinoline - PANH26 | | | | C2 Benzo[a]pyrene- APAH26 | |
| 5-Bromoindole - PANH27 | | | | C1 Chrysene - APAH27 | |
| 9-Chloroacridine- PANH28 | | | | C2 Chrysene - APAH28 | |
| 9-Bromoacridine - PANH29 | | | | C3-Chrysene - APAH29 | |
| 3,6-Dichlorocarbazole - PANH30 | | | | C4-Chrysene - APAH30 | |
| 3,6-Dibromocarbazole- PANH31 | | | | Retene - APAH31 | |
| | | | | 1,2-Benzonapthothiophene - APAH32 | |
| | | | | 2,3-Benzonapthothiophene - APAH33 | |

**Table III 2:** The targeted analytes and their CAS numbers

| Polyaromatic Nitrogen Heterocycles | Polyaromatic Sulfur Heterocycles | Polyaromatic Oxygen Heterocycles | Polycyclic Aromatic Hydrocarbons | Alkylated Polycyclic Aromatic Hydrocarbons | Internal Standards |
|---|---|---|---|---|---|
| Quinoline - 91-22-5 | 4-Methylbenzothiophene - 14315-11-8 | Dibenzofuran – 132-64-9 | Naphthalene- 91-20-3 | 1-Methylnaphthalene – 90-12-0 | Naphthalene-d8 - 1146-65-2 |
| Isoquinoline - 119-65-3 | 2,3-Dimethylbenzothiophene - 4923-91-5 | 2-Methyldibenzofuran - 4265-25-2 | Acenapthene- 83-32-9 | 2-Methylnaphthalene – 91-57-6 | Fluorene-d10 - 81103-79-9 |
| Indole - 120-72-9 | 2,3,4-Trimethylbenzothiophene 70021-44-2 | Benzo[b]naphtho[2,3-d]furan - 243-42-5 | Acenapthylene- 208-96-8 | C2 Naphthalene | Phenanthrene-d10 - 1517-22-2 |
| 2-Methylquinoline – 91-63-4 | 4-Methyldibenzothiophene - 7372-88-5 | Dinaphtho[1,2-b;1',2'-d]furan - 68518-20-7 | Fluorene- 86-73-7 | C3-Naphthalene | Anthracene-d10 - 1719-06-8 |
| 3-Methylisoquinoline - 1125-80-0 | 2,4-Dimethyldibenzothiophene - 31317-18-7 | | Phenanthrene- 1985-01-08 | C4-Naphthalene | Fluoranthene-d10 - 93951-69-0 |
| 2,5-Dimethyl-1-phenylpyrrole - 83-24-9 | 2,4,7-Trimethyldibenzothiophene - 216983-03-8 | | Anthracene - 120-12-7 | C1 Fluorene | Pyrene-d10 - 1718-52-1 |
| 4-Phenylpyridine – 939-23-1 | 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene - 39097-15-9 | | Fluoranthene - 206-44-0 | C2 Fluorene | Chrysene-d12 - 1719-03-5 |
| 2,3-Dimethylindole – 91-55-4 | 6-Methylbenzo [b]naphtho[2,1-d]thiophene - 4567-51-5 | | Pyrene - 129-00-0 | C3 Fluorene | Benz[a]anthracene-d12 - 1718-53-2 |
| Benzo[h]quinoline – 230-27-3 | Dibenzothiophene - 884-12-500 | | Benz[a]anthracene – 56-55-3 | C4 Fluorene | Benzo[b]fluoranthene-d12 - 93951-98-5 |
| 1,2,3,4-Tetrahydrocarbazole - 942-01-8 | C1 Dibenzothiophene | | Chrysene - 218-01-9 | 1,7-Dimethylphenanthrene - 483-87-4 | Benzo[k]fluoranthene-d12 - 93952-01-3 |
| Acridine - 260-94-6 | C2 Dibenzothiophene | | Benzo[a]pyrene - 50-32-8 | 1,8-Dimethylphenanthrene - 7372-87-4 | Benzo[g,h,i]perylene-d12 - 93952-00-2 |
| 9-Methylcarbazole - 1484-12-4 | C3 Dibenzothiophene | | Benzo[b]fluoranthene - 205-99-2 | 1-Methylphenanthrene - 832-69-9 | |
| Carbazole - 86-74-8 | C4 Dibenzothiophene | | Benzo[k]fluoranthene - 207-08-9 | 2,6-Dimethylphenanthrene - 17980-16-4 | |
| 3-Methylbenzo[f]quinoline - 1985-06-03 | 2,3-Dibromothiophene – 3140-93-0 | | Indeno[1,2,3-cd]pyrene - 193-39-5 | 2-Methylphenanthrene - 2531-84-2 | |

| | | | | | |
|---|---|---|---|---|---|
| 2-Methylacridine – 613-15-0 | 5-Chlorobenzothiophene - 20532-33-6 | | Dibenzo[a,h]anthracene - 53-70-3 | 3,6-Dimethylphenanthrene - 1576-67-6 | |
| 3,6-Dimethylcarbazole - 5599-50-8 | 2-Bromobenzothiophene - 5394-13-8 | | Benzo[g,h,i]perylene - 191-24-2 | 3-Methylphenanthrene - 832-71-3 | |
| 3-Tert-butyl-9H-carbazole - 22401-74-7 | 2,3-Dichlorobenzothiophene - 523-97-7 | | | 9/4-Methylphenanthrene - 883-20-5 | |
| 9-Phenylcarbazole - 1150-62-5 | 2,3-Dibromobenzothiophene - 6287-82-7 | | | C2 Phenanthrene | |
| Benz[c]acridine – 225-51-4 | 2-Bromodibenzothiophene - 22439-61-8 | | | C3-Phenanthrene | |
| 11H-Benzo[a]carbazole - 239-01-0 | 2,8-Dibromodibenzohiophene - 31574-87-5 | | | C4- Phenanthrene | |
| Dibenz[c,h]acridine - 224-53-3 | | | | C1 Pyrene | |
| 13H-Dibenzo[a,i]carbazole - 239-64-5 | | | | C2 Pyrene | |
| 2-Chloroquinoline – 612-62-4 | | | | C3 Pyrene | |
| 2-Bromoquinoline - 2005-43-8 | | | | C4 Pyrene | |
| 5-Chloroindole – 17422-32-1 | | | | C1 Benzo[a]pyrene | |
| 3-Bromoisoquinoline - 34784-02-6 | | | | C2 Benzo[a]pyrene | |
| 5-Bromoindole – 10075-50-0 | | | | C1 Chrysene | |
| 9-Chloroacridine- 1207-69-8 | | | | C2 Chrysene | |
| 9-Bromoacridine – 4357-57-7 | | | | C3-Chrysene | |
| 3,6-Dichlorocarbazole - 6825-20-3 | | | | C4-Chrysene | |
| 3,6-Dibromocarbazole- 6825-20-3 | | | | Retene - 483-65-8 | |
| | | | | 1,2-BNT - 205-43-6 | |
| | | | | 2,3- BNT- 243-46-9 | |

**Table III 3:** Average internal standard recoveries for all data calculated with respect to IPIS (Anthracene-d10)

| **ISTD** | **Mean** | **Standard deviation** | **RSD (%)** |
|---|---|---|---|
| Naphthalene-d8 | 34.12 | 13.91 | 40.77 |
| Fluorene-d10 | 69.05 | 13.92 | 20.16 |
| Phenanthrene-d10 | 87.00 | 14.50 | 16.67 |
| Anthracene-d10 | 100.00 | 0.00 | 0.00 |
| Fluoranthene-d10 | 97.61 | 11.97 | 12.26 |
| Pyrene-d11 | 94.07 | 11.72 | 12.46 |
| Chrysene-d12 | 107.98 | 14.37 | 13.31 |
| Benz[a]anthracene-d12 | 99.76 | 13.56 | 13.59 |
| Benzo[b]fluranthene-d12 | 110.13 | 18.60 | 16.89 |
| Benzo[k]fluranthene-d12 | 103.20 | 17.04 | 16.51 |
| Benzo(ghi)perylene-d14 | 102.27 | 18.65 | 18.23 |

**Table III 4:** Average lipid percentage of each species.

| Species Name | **Lipid content** |
|---|---|
| Freshwater drum | 1.1 % |
| Gizzard shad | 18.5 % |
| Yellow perch | 2.0 % |
| Walleye | 7.4 % |

**The protocols followed during sample collection and storage**

The samples were collected between May and October during the 2019 Lake Erie CSMI (Forage Task Group 2023). The intended sample size was at least 5 but with a goal of 10 individuals per basin. Fish were euthanized according to the standard operating procedure of the government agencies conducting the survey (Ontario Ministry of the Natural Resources, Ohio Department of Natural Resources and US Geological Service), which complied with local government animal care procedures and were placed on ice for transportation to the Great Lakes Institute for Environmental Research. Upon dissection, total length and wet weight of each fish was measured and a 1–2 g skinless muscle sample was taken from below the dorsal fin and above the lateral line for stable isotope analysis. The samples were stored under -20 °C prior to the analysis.

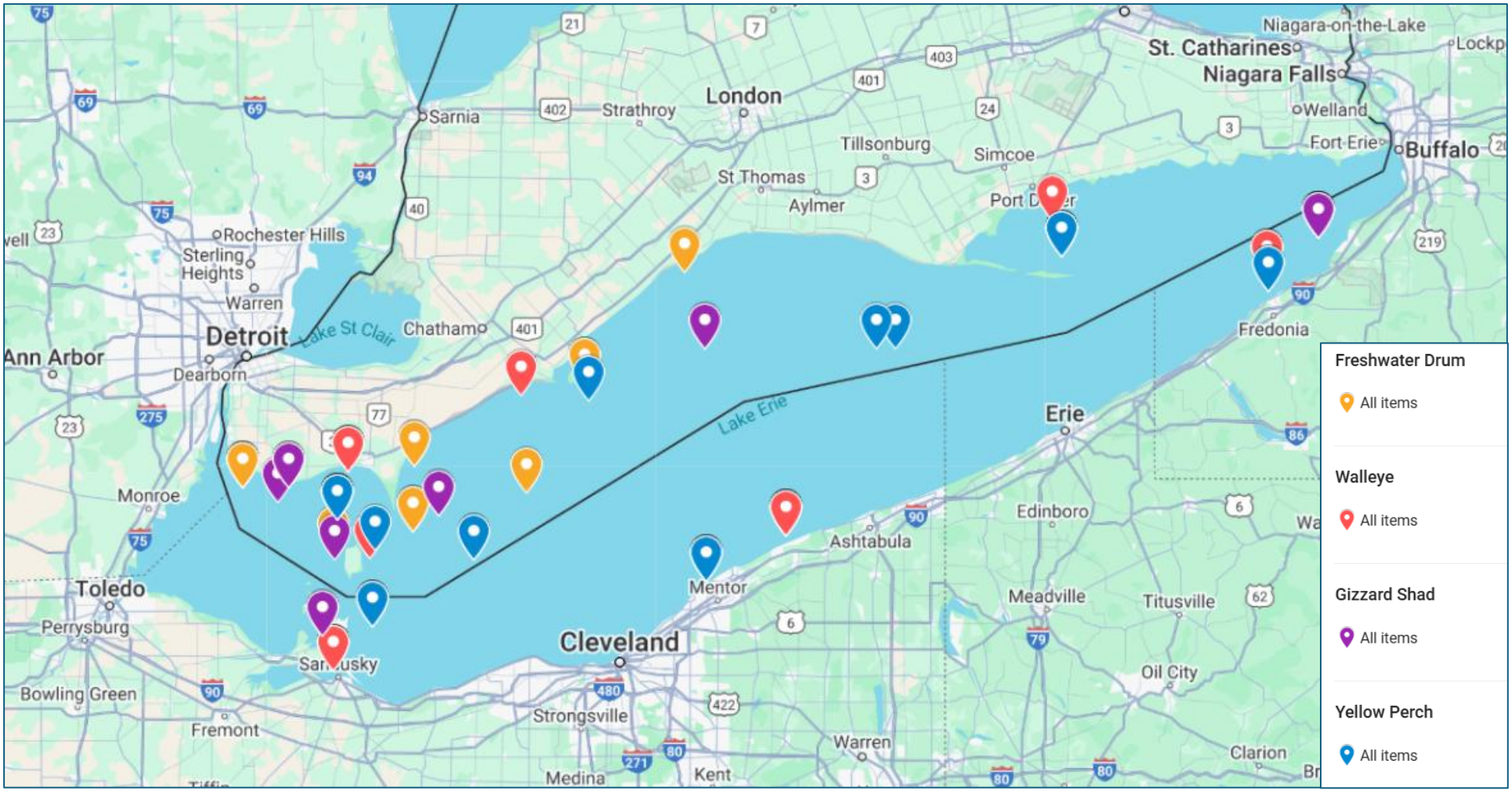



**Figure III 1:** Sample locations from Lake Erie

**Method validation protocols**

Due to the unavailability of commercial standard reference materials store bought wild salmon sample was homogenised, freeze dried and used as the matrix. three concentration levels (25, 125 and 500 pg/µL) were used for the method validation. Six matrix blanks were processed without spiking the HPACs as a quality control measure.

To check accuracy the percent recoveries were calculated by extracting 10 replicants of each concentration level spiked samples. The relative standard deviation (RSD) was calculated for the percent recovery values of each concentration level to assess the repeatability of the method. The intermediate precision was measured using three sets of 125 pg/ µL fortified samples (n=4) that were extracted on three separate days. The RSD were calculated for each sample set and single-factor ANOVA test was used to confirm the intermediate precision.

The detection limits, standard deviations (SD) were calculated based on the 10 pg/ µL spike recoveries (n=10) and blank concentrations (n=6) following the procedure in Eurachem guidelines. Then LOD (3 x SD) and LOQ (10 x SD) values were calculated. The ruggedness of the method was determined by calculating the RSDs of 125 pg/ µL spike recoveries that were extracted in two sets of spike samples. For one set (n=4) the mass of each dispersant (silica gel and 5% deactivated alumina) was deliberately increased by 0.5 g and, the second set was extracted at 140 °C instead of 120 °C (n=4).

**Table III 5:** Performance characteristics of the extraction and purification method used for HPACs

| The Chemical Name | Selected ISTD | % Recovery | | | Repeatability (RSTD%) | | | Intermediate precision at 125 pg/ul (RSTD%) | | LOD (ng/g) | LOQ (ng/g) | Ruggedness % Recovery | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 500 pg/ul | 125 pg/ul | 25 pg/ul | 500 pg/ul | 125 pg/ul | 25 pg/ul | Day 2 | Day 3 | | | Dispersent mass | ASE Tem |
| **HPACs** | | | | | | | | | | | | | |
| Quinoline | Naphthalene d8 | 86% | 102% | 40% | 19.6 | 8.3 | 18.2 | 18.5 | 13.1 | 0.8 | 2.7 | 91% | 94% |
| Isoquinoline | Naphthalene d8 | 82% | 115% | 48% | 27.0 | 10.8 | 27.2 | 23.8 | 18.0 | 1.1 | 3.8 | 74% | 84% |
| Indole | Naphthalene d8 | 90% | 101% | 3% | 14.1 | 17.0 | 33.2 | 23.6 | 14.7 | 2.2 | 7.4 | 144% | 146% |
| 2-Methylquinoline | Naphthalene d8 | 84% | 113% | 56% | 27.2 | 11.1 | 23.8 | 21.8 | 32.8 | 1.1 | 3.6 | 85% | 87% |
| 4-Methylbenzothiophene | Fluorene d10 | 97% | 88% | 95% | 5.7 | 11.5 | 9.1 | 15.7 | 15.7 | 0.8 | 2.8 | 98% | 109% |
| 3-Methylisoquinoline | Naphthalene d8 | 81% | 106% | 64% | 35.2 | 14 | 26.8 | 22.0 | 31.5 | 1.2 | 4.1 | 68% | 77% |
| 2,3dimethylbenzothiophene | Fluorene d10 | 93% | 80% | 85% | 9.5 | 8.6 | 11.4 | 4.9 | 8.5 | 0.6 | 1.9 | 94% | 97% |
| 4-phenylpiridine | Naphthalene d8 | 83% | 121% | 55% | 32.1 | 14.6 | 27.5 | 27.9 | 22.1 | 1.5 | 5.1 | 70% | 87% |
| Dibenzofuran | Fluorene d10 | 107% | 104% | 94% | 5.7 | 3.9 | 5.9 | 3.6 | 16.0 | 0.4 | 1.3 | 137% | 133% |
| 2,3,4 trimethylbenzothiophene | Fluorene d10 | 106% | 97% | 103% | 4.2 | 6.9 | 5.6 | 6.7 | 6.1 | 0.6 | 1.9 | 127% | 125% |
| 2-Methyldibenzofuran | Phenanthrene d 10 | 123% | 121% | 123% | 7.7 | 12.6 | 6.9 | 2.6 | 8.4 | 1.3 | 4.4 | 131% | 128% |
| Dibenzothiophene | Phenanthrene d 10 | 124% | 118% | 128% | 5.6 | 4.8 | 7.0 | 2.9 | 8.0 | 0.5 | 1.7 | 142% | 136% |
| Benzo[h]quinoline | Phenanthrene d 10 | 110% | 106% | 115% | 6.5 | 6.3 | 17.2 | 4.9 | 10.1 | 0.7 | 2.2 | 113% | 129% |
| Acridine | Phenanthrene d 10 | 54% | 68% | 91% | 13.7 | 11.6 | 29.6 | 27.6 | 17.4 | 0.8 | 2.8 | 68% | 87% |
| 9-methylcarbazole | Phenanthrene d 10 | 122% | 120% | 125% | 5.7 | 5.8 | 4.1 | 7.6 | 8.8 | 0.6 | 1.9 | 130% | 137% |
| Carbazole | Phenanthrene d 10 | 130% | 132% | 122% | 6.9 | 4.3 | 9.0 | 5.4 | 20.4 | 0.5 | 1.7 | 117% | 139% |
| 4 methyl dibenzothiophene | Phenanthrene d 10 | 121% | 120% | 138% | 7.9 | 7.7 | 21.4 | 15.3 | 16.6 | 1.0 | 3.4 | 159% | 155% |
| 3-Methylbenzo[f]quinoline | Phenanthrene d 10 | 74% | 92% | 64% | 28.1 | 9.2 | 27.2 | 22.2 | 16.6 | 0.7 | 2.4 | 68% | 72% |
| 2-Methylacridine | Phenanthrene d 10 | 54% | 72% | 60% | 19.7 | 15.4 | 37.9 | 10.1 | 26.3 | 1.4 | 4.6 | 96% | 104% |
| 2,4 dimethyldibenzothiophene | Fluoranthene d11 | 133% | 109% | 107% | 7.0 | 17.7 | 22.2 | 1.4 | 9.9 | 1.8 | 6.1 | 161% | 175% |
| 3,6-Dimethylcarbazole | Fluoranthene d11 | 120% | 115% | 104% | 5.7 | 4.5 | 11.0 | 10.6 | 11.5 | 0.4 | 1.5 | 128% | 148% |
| 2,4,7 trimethyldibenzothiophene | Pyrene d11 | 110% | 105% | 79% | 10.7 | 13.7 | 32.4 | 8.2 | 2.9 | 1.3 | 4.5 | 154% | 157% |
| Benzo[b]naphtho[2,3-d]furan | Pyrene d11 | 132% | 124% | 102% | 15.4 | 12.5 | 39.8 | 5.1 | 3.2 | 1.4 | 4.8 | 151% | 137% |
| 3-tert-butyl-9H-carbazole | Pyrene d11 | 119% | 115% | 114% | 10.6 | 7.7 | 6.7 | 3.2 | 8.5 | 0.7 | 2.5 | 131% | 156% |
| 9-Phenylcarbazole | Benz(a)anthrazene d12 | 117% | 111% | 116% | 19.8 | 12.1 | 6.8 | 1.9 | 6.3 | 1.1 | 3.7 | 151% | 137% |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | Benz(a)anthrazene d12 | 100% | 88% | 113% | 7.0 | 11.7 | 13.3 | 7.1 | 22.1 | 0.9 | 2.9 | 138% | 130% |
| Benz[c]acridine | Benz(a)anthrazene d12 | 111% | 110% | 123% | 11.2 | 7.7 | 20.1 | 14.9 | 11.9 | 0.7 | 2.4 | 136% | 167% |
| 11H-Benzo[a]carbazole | Benz(a)anthrazene d12 | 116% | 112% | 95% | 12.8 | 10.1 | 18.1 | 21.8 | 7.1 | 1.0 | 3.2 | 113% | 166% |
| 7H-Benzo[c]carbazole | Benz(a)anthrazene d12 | 109% | 109% | 83% | 8.5 | 7.4 | 15.9 | 17.1 | 3.3 | 0.7 | 2.3 | 107% | 150% |
| 6-Methylbenzo[b]naphtho[2,1-d]thiophene | Benz(a)anthrazene d12 | 117% | 103% | 113% | 8.1 | 11.7 | 8.4 | 15.2 | 43.0 | 1.0 | 3.3 | 146% | 155% |
| Dinaphtho[1,2-b;1',2'-d]furan | Benzo(ghi)perylene d12 | 116% | 116% | 115% | 9.3 | 12.2 | 8.1 | 6.7 | 3.6 | 1.2 | 3.9 | 124% | 126% |
| Dibenz[c,h]acridine | Benzo(ghi)perylene-d12 | 126% | 120% | 120% | 9.4 | 10.2 | 6.3 | 9.0 | 3.2 | 1.0 | 3.3 | 130% | 136% |
| 13(H)-Dibenzo[a,i]carbazole | Benzo(ghi)perylene-d12 | 123% | 114% | 93% | 8.7 | 8.9 | 12.5 | 13.7 | 7.2 | 0.8 | 2.8 | 109% | 139% |
| **Halo-HPACs** | | | | | | | | | | | | | |
| 2,3-Dibromothiophene | Fluorene d10 | 67% | 87% | 82% | 4.7 | 5.7 | 7.8 | 7.7 | 6.8 | 0.6 | 2.1 | 56% | 54% |
| 5-Chlorobenzothiophene | Fluorene d10 | 80% | 83% | 90% | 2.7 | 7.2 | 3.5 | 6.5 | 3.9 | 0.6 | 2.1 | 78% | 73% |
| 2-chloroquinoline | Fluorene d10 | 91% | 90% | 99% | 2.4 | 6.3 | 2.6 | 5.6 | 3.8 | 0.6 | 1.9 | 89% | 87% |
| 2-Bromobenzothiophene | Fluorene d10 | 75% | 90% | 91% | 3.1 | 4.8 | 4.3 | 7.0 | 2.2 | 0.7 | 2.3 | 79% | 76% |
| 2,3-Dichlorobenzothiophene | Fluorene d10 | 77% | 85% | 90% | 3.5 | 4.5 | 3.9 | 5.2 | 2.2 | 0.5 | 1.6 | 78% | 75% |
| 2-bromoquinoline | Naphthelene d8 | 94% | 95% | 93% | 4.0 | 7.8 | 4.9 | 7.9 | 2.8 | 0.9 | 3.0 | 95% | 94% |
| 5-chloroindole | Fluorene d10 | 79% | 67% | 61% | 4.7 | 13.0 | 10.1 | 19.5 | 4.7 | 1.3 | 4.2 | 77% | 78% |
| 3-bromoisoquinoline | Fluorene d10 | 93% | 96% | 93% | 2.5 | 7.6 | 3.5 | 6.8 | 3.7 | 0.8 | 2.6 | 100% | 101% |
| 5-bromoindol | Naphthalene d8 | 94% | 84% | 67% | 6.0 | 7.3 | 5.9 | 19.4 | 5.2 | 1.4 | 4.7 | 83% | 85% |
| 2,3-Dibromobenzothiophene | Phenanthrene d10 | 74% | 80% | 86% | 4.1 | 7.4 | 4.6 | 6.7 | 6.1 | 0.6 | 2.1 | 72% | 67% |
| 9-chloroacridine | Phenanthrene d10 | 43% | 35% | 48% | 19.2 | 14.1 | 9.1 | 24.6 | 25.3 | 1.7 | 5.6 | 91% | 80% |
| 2-Bromodibenzothiophene | Phenanthrene d10 | 75% | 93% | 100% | 8.2 | 10.4 | 5.7 | 11.2 | 1.3 | 1.2 | 3.9 | 83% | 87% |
| 9-bromoacridine | Phenanthrene d10 | 42% | 37% | 47% | 20.7 | 28.0 | 12.4 | 22.8 | 20.5 | 1.2 | 4.0 | 74% | 74% |
| 3,6-dichloro-9h-carbazole | Pyrene-d11 | 99% | 99% | 94% | 3.1 | 8.1 | 6.6 | 10.5 | 5.8 | 1.1 | 3.8 | 113% | 113% |
| 2,8-Dibromodibenzothiophene | Pyrene-d11 | 99% | 87% | 97% | 3.5 | 6.8 | 4.1 | 7.2 | 4.7 | 0.7 | 2.3 | 101% | 102% |
| 3,6-dibromocarbazole | Chrysene d12 | 86% | 78% | 83% | 4.6 | 9.7 | 5.1 | 9.5 | 7.7 | 0.8 | 2.7 | 96% | 97% |

**Table III 6:** Analyte-species combinations that have significantly different concentrations in fish from eastern and western regions

| PAC | Species | statistic | p.value | Group_Comparison | Higher_Group |
|---|---|---|---|---|---|
| APAHs | Freshwater Drum | -3.73983 | 0.013434 | Different | West Central |
| 1,7-Dimethylphenanthrene | Freshwater Drum | -10.1078 | 0.000539 | Different | West Central |
| 1-Methylphenanthrene | Freshwater Drum | -13.2167 | 5.02E-05 | Different | West Central |
| 2,6-Dimethylphenanthrene | Freshwater Drum | -8.19057 | 0.00121 | Different | West Central |
| 2,6-Dimethylphenanthrene | Walleye | -3.65549 | 0.019757 | Different | West Central |
| 2-Methylphenanthrene | Freshwater Drum | -13.5905 | 0.00017 | Different | West Central |
| 2-Methylphenanthrene | Walleye | -2.94684 | 0.035524 | Different | West Central |
| 3,6-Dimethylphenanthrene | Freshwater Drum | -5.87423 | 0.040741 | Different | West Central |
| 3-Methylphenanthrene | Walleye | -4.70678 | 0.003303 | Different | West Central |
| 9/4-Methylphenanthrene | Freshwater Drum | -6.73183 | 0.002087 | Different | West Central |
| C1 Pyrene | Walleye | -3.18688 | 0.026433 | Different | West Central |
| C2 Pyrene | Freshwater Drum | -3.96378 | 0.016429 | Different | West Central |
| C3 Pyrene | Freshwater Drum | -6.11458 | 0.027171 | Different | West Central |
| C3 Pyrene | Walleye | -8.15232 | 0.014715 | Different | West Central |
| C4 Pyrene | Freshwater Drum | -15.5591 | 9.96E-05 | Different | West Central |
| C2 Chrysene | Freshwater Drum | -3.04146 | 0.035703 | Different | West Central |
| C3 Naphthalene | Walleye | -2.74049 | 0.034012 | Different | West Central |
| C2 Fluorene | Walleye | -5.04045 | 0.004374 | Different | West Central |
| PAHs | Walleye | -3.21054 | 0.029477 | Different | West Central |
| Fluoranthene | Freshwater Drum | -4.33759 | 0.00745 | Different | West Central |
| Fluoranthene | Yellow Perch | 2.49018 | 0.038268 | Different | East Central |
| Benzo[b]fluoranthene | Freshwater Drum | -3.11033 | 0.03586 | Different | West Central |
| Phenanthrene | Freshwater Drum | -3.90467 | 0.011487 | Different | West Central |
| Pyrene | Yellow Perch | -3.86208 | 0.018115 | Different | West Central |
| PANHs | Freshwater Drum | -4.83255 | 0.013857 | Different | West Central |
| 1,2,3,4-Tetrahydrocarbazole | Freshwater Drum | -3.46498 | 0.02281 | Different | West Central |
| Acridine | Freshwater Drum | -2.63267 | 0.048044 | Different | West Central |
| Carbazole | Freshwater Drum | -3.05835 | 0.036527 | Different | West Central |
| 2-Methylacridine | Walleye | -3.09878 | 0.022764 | Different | West Central |
| 11H-Benzo[a]carbazole | Freshwater Drum | -5.41455 | 0.004694 | Different | West Central |
| Dibenz[c,h]acridine | Yellow Perch | 4.799753 | 0.001857 | Different | East Central |
| 5-Chloroindole | Freshwater Drum | -9.21103 | 0.00028 | Different | West Central |
| 5-Chloroindole | Yellow Perch | -3.26416 | 0.014958 | Different | West Central |
| 5-Bromoindole | Yellow Perch | -4.37878 | 0.004923 | Different | West Central |
| 9-Chloroacridine | Yellow Perch | -2.89733 | 0.019983 | Different | West Central |

| | | | | | |
|---|---|---|---|---|---|
| 3,6-Dichlorocarbazole | Walleye | 3.040568 | 0.022792 | Different | East Central |
| 3,6-Dichlorocarbazole | Yellow Perch | -6.54531 | 0.000366 | Different | West Central |
| 2-Methylquinoline | Yellow Perch | 3.874983 | 0.005623 | Different | East Central |
| Dibenzofuran | Yellow Perch | -3.91435 | 0.01733 | Different | West Central |
| 2-Methyldibenzofuran | Yellow Perch | 2.526216 | 0.038929 | Different | East Central |
| Benzo[b]naphtho[2,3-d]furan | Walleye | -2.79866 | 0.031221 | Different | West Central |
| PASHs | Freshwater Drum | -3.27296 | 0.022385 | Different | West Central |
| C1 Dibenzothiophene | Freshwater Drum | -35.7461 | 3.66E-06 | Different | West Central |
| C1 Dibenzothiophene | Walleye | -4.11648 | 0.008387 | Different | West Central |
| C2 Dibenzothiophene | Freshwater Drum | -3.43681 | 0.026371 | Different | West Central |
| C2 Dibenzothiophene | Walleye | -3.9319 | 0.010879 | Different | West Central |
| C3 Dibenzothiophene | Walleye | -2.70734 | 0.035419 | Different | West Central |
| C4 Dibenzothiophene | Freshwater Drum | 3.599441 | 0.022591 | Different | East Central |
| 2-Bromobenzothiophene | Freshwater Drum | -6.47985 | 0.005586 | Different | West Central |
| 2,3-Dibromobenzothiophene | Yellow Perch | -3.64339 | 0.019177 | Different | West Central |
| 2,3-Dimethylbenzothiophene | Yellow Perch | 3.20302 | 0.023132 | Different | East Central |
| Dibenzothiophene | Freshwater Drum | -2.82956 | 0.047546 | Different | West Central |
| 2,4,7-Trimethyldibenzothiophene | Freshwater Drum | -8.12726 | 0.024247 | Different | West Central |
| 2,4,7-Trimethyldibenzothiophene | Yellow Perch | 2.984839 | 0.017494 | Different | East Central |
| 6-Methylbenzo [b]naphtho[2,1-d]thiophene | Freshwater Drum | -3.21492 | 0.032437 | Different | West Central |




**Figure III 2:** PANH concentrations detected in each species (FD: freshwater drum, GS; gizzard shad, YP: yellow perch, WE: walleye, WE: western, WC: west central, EC: east central, ES: eastern)

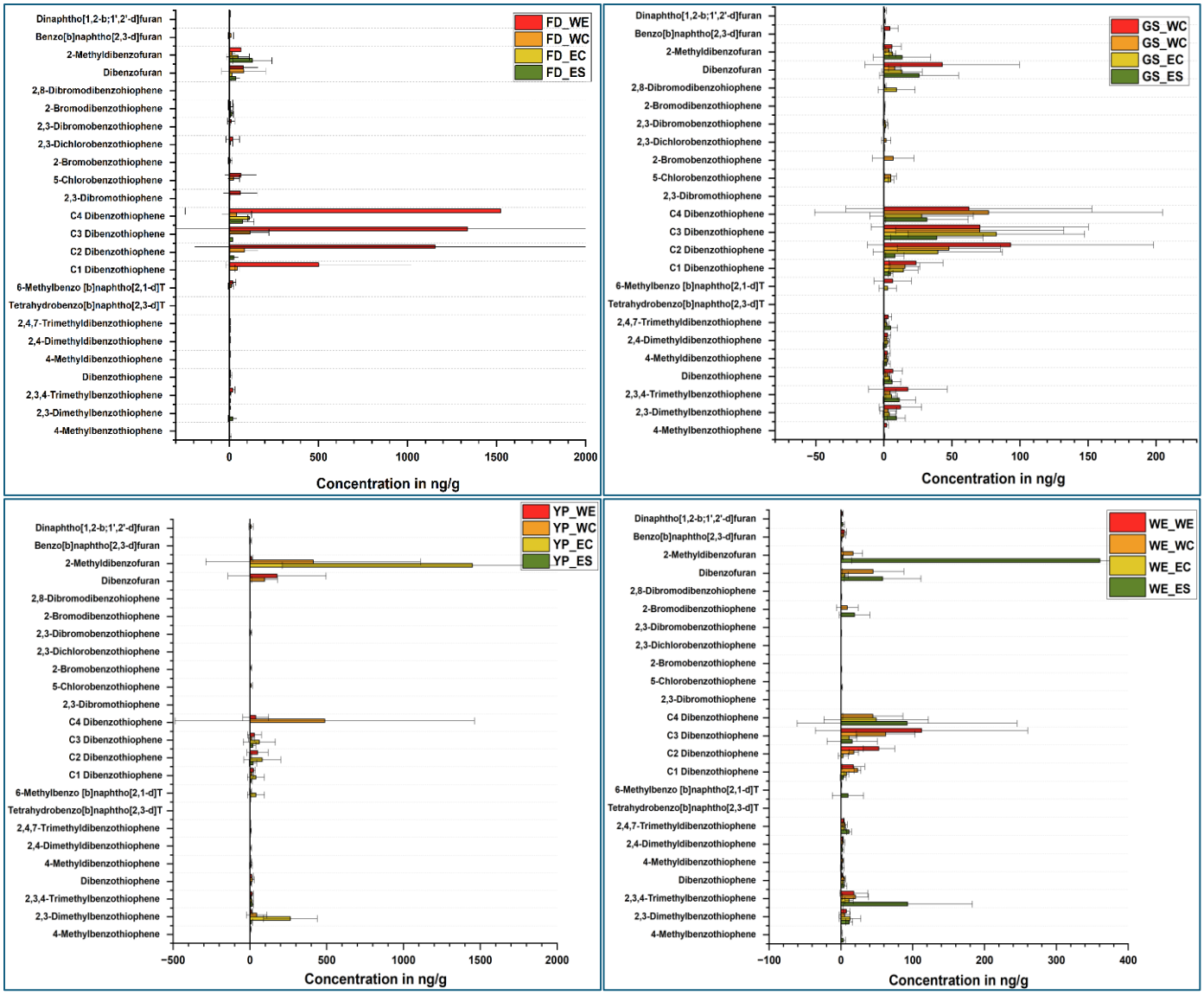



**Figure III 3:** PASH and PAOH concentrations detected in each species (FD: freshwater drum, GS; gizzard shad, YP: yellow perch, WE: walleye, WE: western, WC: west central, EC: east central, ES: eastern).

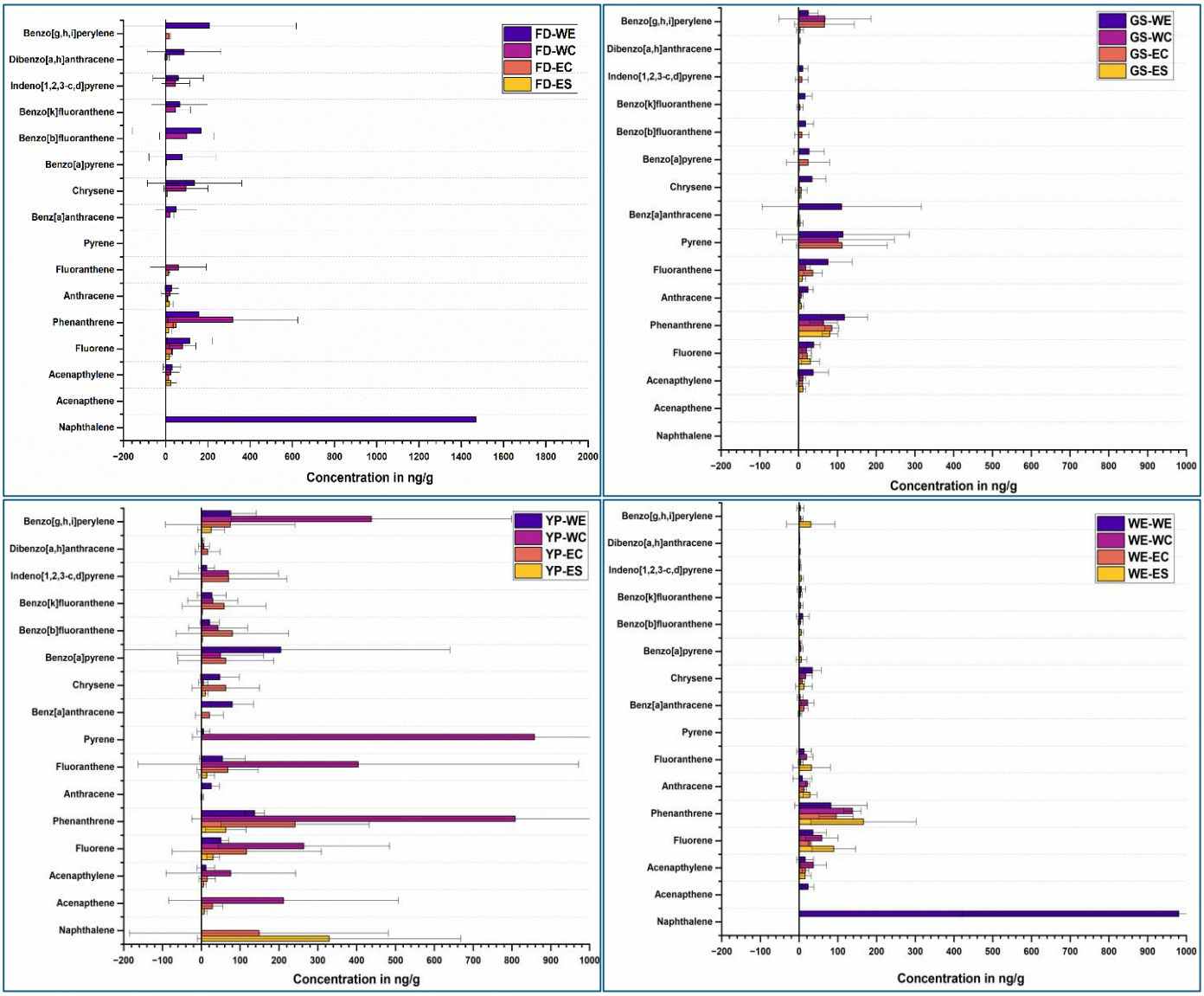



**Figure III 4:** Lipid corrected PAH concentrations detected in each species (FD: freshwater drum, GS; gizzard shad, YP: yellow perch, WE: walleye, WE: western, WC: west central, EC: east central, ES: east

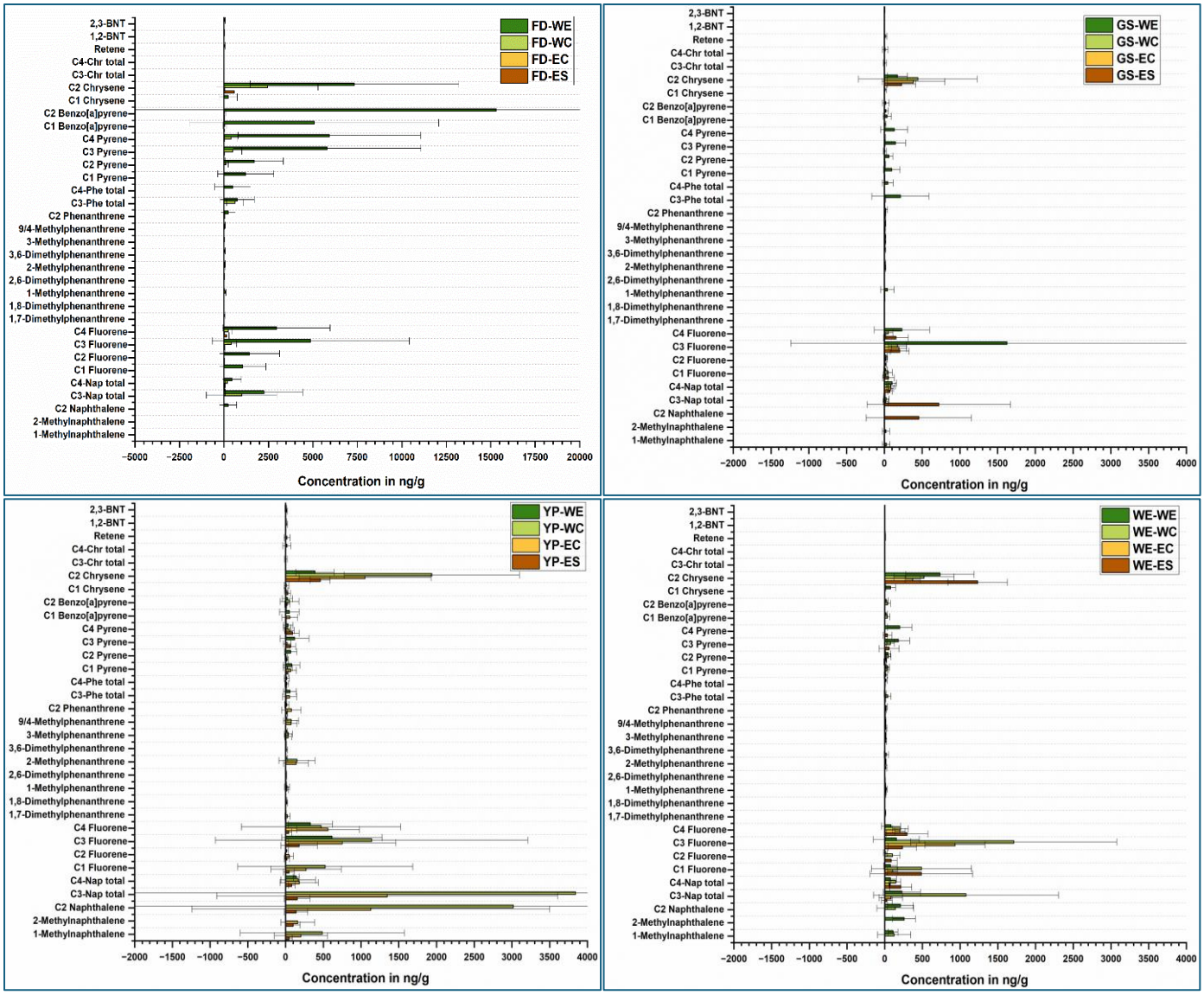



**Figure III 5:** Lipid corrected APAH concentrations detected in each species (FD: freshwater drum, GS; gizzard shad, YP: yellow perch, WE: walleye, WE: western, WC: west central, EC: east central, ES: eastern).

**Table III 7:** The selected analytes for UMAP and PCA and their loading scores for the first five principal components

| Analyte | Dim.1 | Dim.2 | Dim.3 | Dim.4 | Dim.5 |
|---|---|---|---|---|---|
| Quinoline | 0.564 | -0.533 | 0.254 | -0.194 | -0.131 |
| 9-Methylcarbazole | 0.422 | -0.180 | 0.459 | 0.324 | 0.010 |
| Carbazole | 0.564 | -0.258 | 0.067 | -0.084 | -0.292 |
| 2-Methylacridine | 0.680 | 0.160 | 0.079 | 0.287 | 0.218 |
| 3,6-Dimethylcarbazole | -0.152 | 0.485 | 0.273 | 0.102 | -0.216 |
| Isoquinoline | 0.391 | -0.615 | 0.153 | 0.023 | -0.112 |
| 11H-Benzo[a]carbazole | 0.222 | 0.090 | 0.092 | 0.718 | 0.273 |
| Dibenz[c,h]acridine | 0.366 | 0.033 | 0.093 | 0.632 | 0.160 |
| 5-Chloroindole | -0.784 | 0.224 | 0.346 | -0.081 | 0.036 |
| 3-Bromoisoquinoline | 0.516 | -0.072 | 0.324 | -0.498 | 0.363 |
| 2-Methylquinoline | 0.465 | -0.280 | -0.023 | -0.035 | 0.290 |
| 3-Methylisoquinoline | 0.534 | -0.529 | -0.059 | 0.078 | 0.093 |
| Dibenzofuran | -0.046 | -0.002 | -0.757 | 0.338 | 0.170 |
| C1 Dibenzothiophene | 0.178 | 0.458 | 0.321 | 0.082 | -0.306 |
| 2,3-Dibromobenzothiophene | 0.018 | 0.246 | -0.071 | -0.471 | 0.619 |
| 2,3-Dimethylbenzothiophene | 0.336 | 0.176 | -0.770 | 0.166 | -0.125 |
| 2,3,4-Trimethylbenzothiophene | 0.474 | -0.034 | -0.374 | -0.317 | -0.263 |
| Dibenzothiophene | 0.551 | 0.541 | -0.114 | -0.256 | 0.377 |
| 4-Methyldibenzothiophene | 0.578 | 0.520 | 0.346 | 0.100 | -0.004 |
| 2,4-Dimethyldibenzothiophene | 0.629 | 0.637 | 0.037 | -0.036 | -0.212 |
| 2,4,7-Trimethyldibenzothiophene | 0.690 | 0.278 | -0.246 | -0.262 | -0.318 |

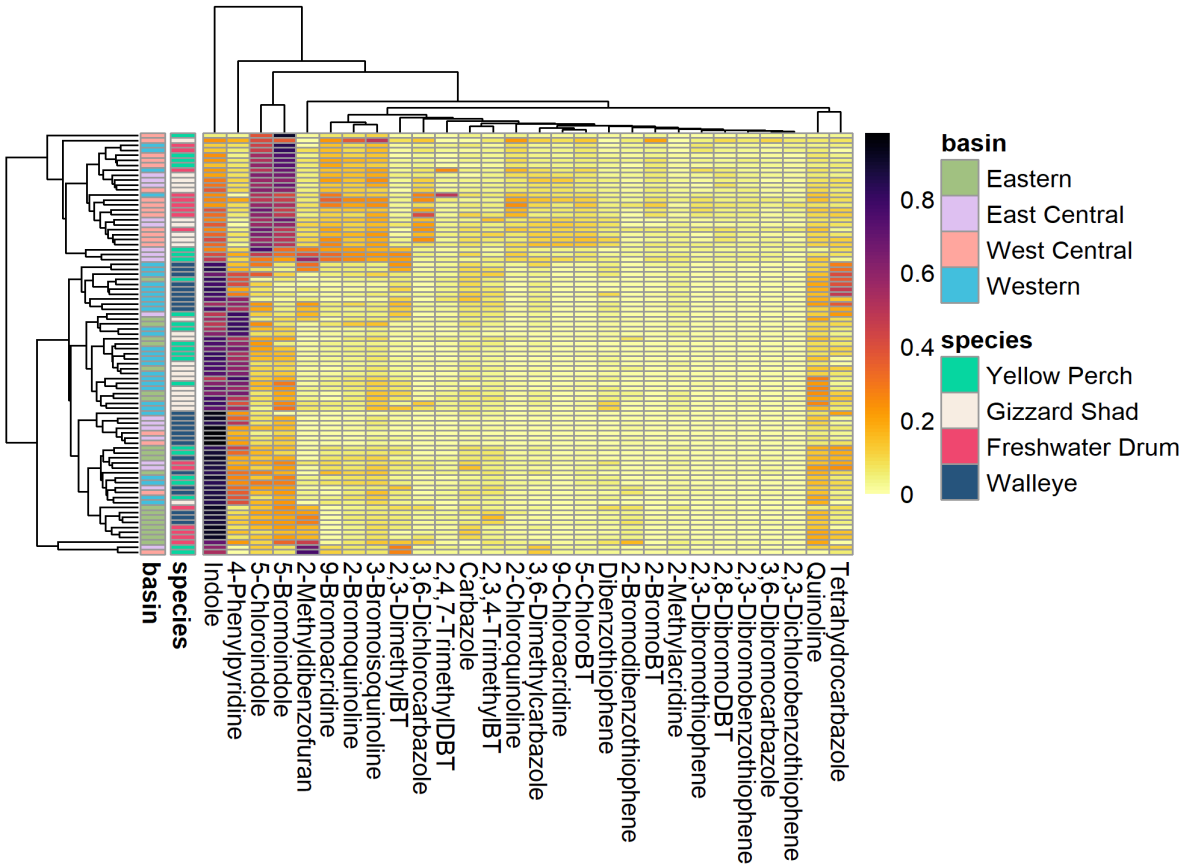



**Figure III 6:** Heatmaps generated from hierarchical cluster analysis results of all four species collected from the four regions of Lake Erie

# APPENDIX IV. ADDITIONAL METRICS FROM CHAPTER 6: HABITAT AND DIET AFFECT BIOACCUMULATION OF HOMOCYCLIC AND HETEROCYCLIC AROMATIC COMPOUNDS IN LAKE ERIE FISH

**Table IV 1:** The targeted analytes and their reference codes used in this study

| Polyaromatic Nitrogen Heterocycles | Polyaromatic Sulfur Heterocycles | Polyaromatic Oxygen Heterocycles | Polycyclic Aromatic Hydrocarbons | Alkylated Polycyclic Aromatic Hydrocarbons | Internal Standards |
|---|---|---|---|---|---|
| Quinoline - PANH1 | 4-Methylbenzothiophene - PASH1 | Dibenzofuran - PAOH1 | Naphthalene- PAH1 | 1-Methylnaphthalene - APAH1 | Naphthalene-d8 |
| Isoquinoline - PANH2 | 2,3-Dimethylbenzothiophene - PASH2 | 2-Methyldibenzofuran - PAOH2 | Acenapthene- PAH2 | 2-Methylnaphthalene - APAH2 | Fluorene-d10 |
| Indole - PANH3 | 2,3,4-Trimethylbenzothiophene - PASH3 | Benzo[b]naphtho[2,3-d]furan - PAOH3 | Acenapthylene- PAH3 | C2 Naphthalene - APAH3 | Phenanthrene-d10 |
| 2-Methylquinoline - PANH4 | 4-Methyldibenzothiophene - PASH4 | Dinaphtho[1,2-b;1',2'-d]furan - PAOH4 | Fluorene- PAH4 | C3-Naphthalene - APAH4 | Anthracene-d10 |
| 3-Methylisoquinoline - PANH5 | 2,4-Dimethyldibenzothiophene - PASH5 | | Phenanthrene- PAH5 | C4-Naphthalene - APAH5 | Fluoranthene-d10 |
| 2,5-Dimethyl-1-phenylpyrrole - PANH6 | 2,4,7-Trimethyldibenzothiophene - PASH6 | | Anthracene- PAH6 | C1 Fluorene- APAH6 | Pyrene-d11 |
| 4-Phenylpyridine - PANH7 | 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene - PASH7 | | Fluoranthene- PAH7 | C2 Fluorene- APAH7 | Chrysene-d12 |
| 2,3-Dimethylindole - PANH8 | 6-Methylbenzo [b]naphtho[2,1-d]thiophene - PASH8 | | Pyrene- PAH8 | C3 Fluorene- APAH8 | Benz[a]anthracene-d12 |
| Benzo[h]quinoline - PANH9 | Dibenzothiophene- PASH9 | | Benz[a]anthracene- PAH9 | C4 Fluorene- APAH9 | Benzo[b]fluranthene-d12 |
| 1,2,3,4-Tetrahydrocarbazole - PANH10 | C1 Dibenzothiophene- PASH10 | | Chrysene- PAH10 | 1,7-Dimethylphenanthrene - APAH10 | Benzo[k]fluranthene-d12 |
| Acridine - PANH11 | C2 Dibenzothiophene- PASH11 | | Benzo[a]pyrene- PAH11 | 1,8-Dimethylphenanthrene - APAH11 | Benzo(ghi)perylene-d14 |
| 9-Methylcarbazole - PANH12 | C3 Dibenzothiophene- PASH12 | | Benzo[b]fluoranthene- PAH12 | 1-Methylphenanthrene - APAH12 | |
| Carbazole - PANH13 | C4 Dibenzothiophene- PASH13 | | Benzo[k]fluoranthene- PAH13 | 2,6-Dimethylphenanthrene - APAH13 | |
| 3-Methylbenzo[f]quinoline - PANH14 | 2,3-Dibromothiophene - PASH14 | | Indeno[1,2,3-c,d]pyrene- PAH14 | 2-Methylphenanthrene - APAH14 | |
| 2-Methylacridine - PANH15 | 5-Chlorobenzothiophene - PASH15 | | Dibenzo[a,h]anthracene- PAH15 | 3,6-Dimethylphenanthrene - APAH15 | |
| 3,6-Dimethylcarbazole - PANH16 | 2-Bromobenzothiophene - PASH16 | | Benzo[g,h,i]perylene- PAH16 | 3-Methylphenanthrene - APAH16 | |

| | | | | | |
|---|---|---|---|---|---|
| 3-Tert-butyl-9H-carbazole - PANH17 | 2,3-Dichlorobenzothiophene - PASH17 | | | 9/4-Methylphenanthrene - APAH17 | |
| 9-Phenylcarbazole - PANH18 | 2,3-Dibromobenzothiophene - PASH18 | | | C2 Phenanthrene - APAH18 | |
| Benz[c]acridine - PANH19 | 2-Bromodibenzothiophene - PASH19 | | | C3-Phenanthrene - APAH19 | |
| 11H-Benzo[a]carbazole - PANH20 | 2,8-Dibromodibenzohiophene - PASH20 | | | C4- Phenanthrene - APAH20 | |
| Dibenz[c,h]acridine - PANH21 | | | | C1 Pyrene- APAH21 | |
| 13H-Dibenzo[a,i]carbazole - PANH22 | | | | C2 Pyrene- APAH22 | |
| 2-Chloroquinoline - PANH23 | | | | C3 Pyrene- APAH23 | |
| 2-Bromoquinoline - PANH24 | | | | C4 Pyrene- APAH24 | |
| 5-Chloroindole - PANH25 | | | | C1 Benzo[a]pyrene- APAH25 | |
| 3-Bromoisoquinoline - PANH26 | | | | C2 Benzo[a]pyrene- APAH26 | |
| 5-Bromoindole - PANH27 | | | | C1 Chrysene - APAH27 | |
| 9-Chloroacridine- PANH28 | | | | C2 Chrysene - APAH28 | |
| 9-Bromoacridine - PANH29 | | | | C3-Chrysene - APAH29 | |
| 3,6-Dichlorocarbazole - PANH30 | | | | C4-Chrysene - APAH30 | |
| 3,6-Dibromocarbazole- PANH31 | | | | Retene - APAH31 | |
| | | | | 1,2-Benzonapthothiophene - APAH32 | |
| | | | | 2,3-Benzonapthothiophene - APAH33 | |

**Table IV 2:** The targeted analytes and their CAS numbers

| Polyaromatic Nitrogen Heterocycles | Polyaromatic Sulfur Heterocycles | Polyaromatic Oxygen Heterocycles | Polycyclic Aromatic Hydrocarbons | Alkylated Polycyclic Aromatic Hydrocarbons | Internal Standards |
|---|---|---|---|---|---|
| Quinoline - 91-22-5 | 4-Methylbenzothiophene - 14315-11-8 | Dibenzofuran – 132-64-9 | Naphthalene- 91-20-3 | 1-Methylnaphthalene – 90-12-0 | Naphthalene-d8 - 1146-65-2 |
| Isoquinoline - 119-65-3 | 2,3-Dimethylbenzothiophene - 4923-91-5 | 2-Methyldibenzofuran - 4265-25-2 | Acenapthene- 83-32-9 | 2-Methylnaphthalene – 91-57-6 | Fluorene-d10 - 81103-79-9 |
| Indole - 120-72-9 | 2,3,4-Trimethylbenzothiophene 70021-44-2 | Benzo[b]naphtho[2,3-d]furan - 243-42-5 | Acenapthylene- 208-96-8 | C2 Naphthalene | Phenanthrene-d10 - 1517-22-2 |
| 2-Methylquinoline – 91-63-4 | 4-Methyldibenzothiophene - 7372-88-5 | Dinaphtho[1,2-b;1',2'-d]furan - 68518-20-7 | Fluorene- 86-73-7 | C3-Naphthalene | Anthracene-d10 - 1719-06-8 |
| 3-Methylisoquinoline - 1125-80-0 | 2,4-Dimethyldibenzothiophene - 31317-18-7 | | Phenanthrene- 1985-01-08 | C4-Naphthalene | Fluoranthene-d10 - 93951-69-0 |
| 2,5-Dimethyl-1-phenylpyrrole - 83-24-9 | 2,4,7-Trimethyldibenzothiophene - 216983-03-8 | | Anthracene - 120-12-7 | C1 Fluorene | Pyrene-d10 - 1718-52-1 |
| 4-Phenylpyridine – 939-23-1 | 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene - 39097-15-9 | | Fluoranthene - 206-44-0 | C2 Fluorene | Chrysene-d12 - 1719-03-5 |
| 2,3-Dimethylindole – 91-55-4 | 6-Methylbenzo [b]naphtho[2,1-d]thiophene - 4567-51-5 | | Pyrene - 129-00-0 | C3 Fluorene | Benz[a]anthracene-d12 - 1718-53-2 |
| Benzo[h]quinoline – 230-27-3 | Dibenzothiophene - 884-12-500 | | Benz[a]anthracene – 56-55-3 | C4 Fluorene | Benzo[b]fluoranthene-d12 - 93951-98-5 |
| 1,2,3,4-Tetrahydrocarbazole - 942-01-8 | C1 Dibenzothiophene | | Chrysene - 218-01-9 | 1,7-Dimethylphenanthrene - 483-87-4 | Benzo[k]fluoranthene-d12 - 93952-01-3 |
| Acridine - 260-94-6 | C2 Dibenzothiophene | | Benzo[a]pyrene - 50-32-8 | 1,8-Dimethylphenanthrene - 7372-87-4 | Benzo[g,h,i]perylene-d12 - 93952-00-2 |
| 9-Methylcarbazole - 1484-12-4 | C3 Dibenzothiophene | | Benzo[b]fluoranthene - 205-99-2 | 1-Methylphenanthrene - 832-69-9 | |
| Carbazole - 86-74-8 | C4 Dibenzothiophene | | Benzo[k]fluoranthene - 207-08-9 | 2,6-Dimethylphenanthrene - 17980-16-4 | |
| 3-Methylbenzo[f]quinoline - 1985-06-03 | 2,3-Dibromothiophene – 3140-93-0 | | Indeno[1,2,3-cd]pyrene - 193-39-5 | 2-Methylphenanthrene - 2531-84-2 | |

| | | | | | |
|---|---|---|---|---|---|
| 2-Methylacridine – 613-15-0 | 5-Chlorobenzothiophene - 20532-33-6 | | Dibenzo[a,h]anthracene - 53-70-3 | 3,6-Dimethylphenanthrene - 1576-67-6 | |
| 3,6-Dimethylcarbazole - 5599-50-8 | 2-Bromobenzothiophene - 5394-13-8 | | Benzo[g,h,i]perylene - 191-24-2 | 3-Methylphenanthrene - 832-71-3 | |
| 3-Tert-butyl-9H-carbazole - 22401-74-7 | 2,3-Dichlorobenzothiophene - 523-97-7 | | | 9/4-Methylphenanthrene - 883-20-5 | |
| 9-Phenylcarbazole - 1150-62-5 | 2,3-Dibromobenzothiophene - 6287-82-7 | | | C2 Phenanthrene | |
| Benz[c]acridine – 225-51-4 | 2-Bromodibenzothiophene - 22439-61-8 | | | C3-Phenanthrene | |
| 11H-Benzo[a]carbazole - 239-01-0 | 2,8-Dibromodibenzohiophene - 31574-87-5 | | | C4- Phenanthrene | |
| Dibenz[c,h]acridine - 224-53-3 | | | | C1 Pyrene | |
| 13H-Dibenzo[a,i]carbazole - 239-64-5 | | | | C2 Pyrene | |
| 2-Chloroquinoline – 612-62-4 | | | | C3 Pyrene | |
| 2-Bromoquinoline - 2005-43-8 | | | | C4 Pyrene | |
| 5-Chloroindole – 17422-32-1 | | | | C1 Benzo[a]pyrene | |
| 3-Bromoisoquinoline - 34784-02-6 | | | | C2 Benzo[a]pyrene | |
| 5-Bromoindole – 10075-50-0 | | | | C1 Chrysene | |
| 9-Chloroacridine- 1207-69-8 | | | | C2 Chrysene | |
| 9-Bromoacridine – 4357-57-7 | | | | C3-Chrysene | |
| 3,6-Dichlorocarbazole - 6825-20-3 | | | | C4-Chrysene | |
| 3,6-Dibromocarbazole- 6825-20-3 | | | | Retene - 483-65-8 | |
| | | | | 1,2-BNT - 205-43-6 | |
| | | | | 2,3- BNT- 243-46-9 | |

**Table IV 3:** The sampling information and isotope data for the fish samples

| Species Code | TL (mm) | Wet Weight (g) | $d^{13}C$ | perC | $d^{15}N$ | perN | CN | $d^{34}S$ | perS | Site_Depth | LAT1 | LONG1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RG1 | 46 | 1 | -23.53 | 44.43 | 15.29 | 13.50 | 3.29 | 1.67 | 1.00 | 19.00 | 41.778 | -82.2270 |
| RG2 | 37 | 0.6 | -23.90 | 44.26 | 14.67 | 13.77 | 3.22 | -0.25 | 1.25 | 23.00 | 41.959 | -82.0362 |
| RG3 | 49 | 1.6 | -22.77 | 44.30 | 15.52 | 13.73 | 3.23 | 2.15 | 1.34 | 21.00 | 42.352 | -80.7025 |
| RG4 | 48 | 1.3 | -25.20 | 45.11 | 13.75 | 13.06 | 3.45 | -0.23 | 0.91 | 7.80 | 41.935 | -82.5217 |
| FD1 | 50 | 0.9 | -23.26 | 44.27 | 13.08 | 14.07 | 3.15 | 1.39 | 1.27 | 11.00 | 41.854 | -82.4447 |
| FD2 | 351 | 483 | -21.08 | 44.49 | 12.07 | 13.69 | 3.25 | -0.10 | 1.05 | 11.00 | 41.854 | -82.4447 |
| FD3 | 233 | 122.2 | -24.90 | 49.28 | 12.62 | 10.44 | 4.72 | 0.06 | 0.78 | 11.00 | 41.854 | -82.4447 |
| FD4 | 192 | 72.2 | -22.59 | 44.44 | 14.61 | 13.69 | 3.25 | 2.67 | 1.44 | 11.00 | 41.854 | -82.4447 |
| WS | 193 | 82 | -21.68 | 42.03 | 10.57 | 13.02 | 3.23 | 0.08 | 0.97 | 9.10 | 41.448 | -82.6225 |
| WS | 231 | 149.4 | -21.48 | 45.52 | 9.75 | 13.14 | 3.47 | 2.76 | 1.03 | 11.00 | 41.854 | -82.4447 |
| WS | 204 | 88 | -19.87 | 42.65 | 9.33 | 12.93 | 3.30 | 0.79 | 0.99 | 9.10 | 41.448 | -82.6225 |
| WS | 415 | 717 | -22.55 | 43.17 | 12.59 | 13.45 | 3.21 | 2.68 | 1.20 | 9.60 | 42.002 | -82.7533 |
| WS | 546 | 1688 | -25.67 | 51.80 | 14.50 | 10.59 | 4.89 | 1.67 | 1.17 | 11.00 | 42.599 | -80.1038 |
| ES | 94 | 7 | -22.22 | 43.52 | 12.46 | 13.65 | 3.19 | 3.10 | 1.09 | 9.60 | 42.002 | -82.7533 |
| ES | 54 | 0.9 | -22.61 | 44.21 | 15.57 | 13.57 | 3.26 | -1.02 | 1.26 | 12.80 | 42.034 | -82.4397 |
| ES | 84 | 5 | -22.70 | 48.15 | 15.53 | 12.48 | 3.86 | 0.36 | 1.20 | 11.00 | 42.599 | -80.1038 |
| ES | 95 | 6 | -23.11 | 48.69 | 16.13 | 12.24 | 3.98 | 0.74 | 1.27 | 11.00 | 42.599 | -80.1038 |
| ES | 93 | 6 | -22.92 | 48.06 | 11.45 | 12.39 | 3.88 | -2.23 | 0.90 | 11.00 | 41.767 | -82.9693 |
| ES | 91 | 6 | -22.13 | 44.73 | 14.66 | 13.90 | 3.22 | 1.94 | 1.33 | 13.00 | 42.698 | -80.1372 |
| ES | 58 | 1.3 | -20.09 | 45.41 | 15.72 | 14.20 | 3.20 | 2.66 | 1.23 | 13.00 | 42.698 | -80.1372 |
| SS | 121 | 14.6 | -21.74 | 43.64 | 12.38 | 14.00 | 3.12 | 1.98 | 1.27 | NA | NA | NA |
| SS | 130 | 20.5 | -21.39 | 43.05 | 15.04 | 13.87 | 3.10 | 1.56 | 1.35 | NA | NA | NA |
| SS | 106 | 10.7 | -21.69 | 44.59 | 15.29 | 14.08 | 3.17 | 2.99 | 1.26 | NA | NA | NA |
| SS | 116 | 14.8 | -23.66 | 44.38 | 12.80 | 13.95 | 3.18 | 3.16 | 1.10 | NA | NA | NA |
| SS | 128 | 20.1 | -22.04 | 44.50 | 15.03 | 14.47 | 3.08 | 2.30 | 1.23 | NA | NA | NA |
| SS | 121 | 15 | -21.58 | 43.87 | 15.23 | 13.90 | 3.16 | 1.80 | 1.25 | NA | NA | NA |
| SS | 122 | 15.1 | -21.92 | 43.89 | 14.74 | 13.98 | 3.14 | 2.16 | 1.21 | NA | NA | NA |
| SS | 124 | 19.5 | -21.50 | 43.29 | 14.98 | 13.40 | 3.23 | 2.26 | 1.25 | NA | NA | NA |
| WE | 235 | 94 | -22.41 | 44.17 | 16.08 | 14.39 | 3.07 | 2.28 | 1.33 | 10.60 | 41.805 | -82.5855 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| WE | 158 | 29 | -21.35 | 44.68 | 17.21 | 14.26 | 3.13 | 2.89 | 2.13 | 4.90 | 41.488 | -82.6759 |
| WE | 97 | 7 | -22.38 | 45.62 | 15.70 | 14.07 | 3.24 | 2.68 | 1.34 | 8.00 | 41.780 | -82.7320 |
| WE | 215 | 73 | -22.30 | 45.47 | 16.99 | 14.39 | 3.16 | 3.57 | 1.72 | 4.90 | 41.488 | -82.6759 |
| WE | 68 | 2.4 | -21.68 | 45.99 | 16.97 | 14.71 | 3.13 | 2.52 | 1.24 | 4.30 | 41.557 | -82.7986 |
| WE | 84 | 5 | -23.32 | 46.39 | 14.61 | 14.71 | 3.15 | 3.36 | 1.01 | 8.50 | 41.972 | -82.5452 |
| WE | 248 | 114 | -23.34 | 44.59 | 10.54 | 13.74 | 3.25 | 4.32 | 1.02 | 21.64 | 42.650 | -79.1770 |
| WE | 251 | 112 | -23.82 | 43.78 | 12.39 | 12.60 | 3.48 | 3.99 | 1.06 | 21.64 | 42.650 | -79.1770 |
| WE | 299 | 197 | -24.00 | 44.48 | 10.47 | 13.84 | 3.21 | 4.43 | 1.09 | 21.64 | 42.650 | -79.1770 |
| WE | 250 | 111 | -23.48 | 43.32 | 11.66 | 12.85 | 3.37 | 3.19 | 1.15 | 21.64 | 42.650 | -79.1770 |
| WB | 92 | 9 | -21.95 | 45.99 | 14.44 | 14.67 | 3.14 | 2.95 | 1.15 | 14.30 | 41.598 | -82.5943 |
| WB | NA | NA | -23.71 | 45.89 | 13.84 | 14.61 | 3.14 | NA | NA | 12.50 | 41.845 | -81.1006 |
| WB | 185 | 69.9 | -23.51 | 47.07 | 12.43 | 14.42 | 3.26 | NA | NA | 8.90 | 42.020 | -82.7043 |
| WB | 101 | 12.4 | -23.35 | 47.75 | 13.65 | 13.63 | 3.50 | NA | NA | 8.90 | 42.020 | -82.7043 |
| WB | 168 | 50.6 | -22.98 | 47.44 | 13.25 | 14.22 | 3.34 | NA | NA | 8.90 | 42.020 | -82.7043 |
| WB | 124 | 20.8 | -23.27 | 47.62 | 13.18 | 14.23 | 3.35 | NA | NA | 8.90 | 42.020 | -82.7043 |
| WB | 329 | 440 | -23.27 | 46.72 | 12.64 | 13.93 | 3.36 | NA | NA | 8.90 | 42.020 | -82.7043 |
| WB | 148 | 78 | -23.02 | 47.57 | 13.18 | 14.24 | 3.34 | NA | NA | 8.90 | 42.020 | -82.7043 |
| YP | 61 | 2 | -22.25 | 45.48 | 15.31 | 14.38 | 3.16 | 3.64 | 1.11 | 26.20 | 42.548 | -79.3593 |
| YP | 179 | 74 | -21.98 | 44.50 | 15.71 | 14.13 | 3.15 | 3.14 | 1.17 | 19.81 | 42.507 | -79.3582 |
| YP | 170 | 54 | -23.12 | 44.14 | 14.96 | 14.00 | 3.15 | 4.27 | 1.08 | 19.81 | 42.507 | -79.3582 |
| YP | 162 | 66 | -23.24 | 45.42 | 12.16 | 14.16 | 3.21 | 3.74 | 1.04 | 19.81 | 42.507 | -79.3582 |
| YP | 173 | 59 | -24.07 | 46.12 | 12.54 | 14.03 | 3.29 | 3.18 | 1.15 | 25.90 | 42.640 | -79.3024 |
| YP | 59 | 2 | -23.68 | 44.91 | 12.28 | 13.59 | 3.30 | 3.61 | 1.08 | 25.90 | 42.640 | -79.3024 |
| YP | 62 | 2 | -23.85 | 43.17 | 13.05 | 13.78 | 3.13 | 2.96 | 1.22 | 12.50 | 41.845 | -81.1006 |
| YP | 66 | 3 | -23.77 | 43.81 | 14.05 | 13.91 | 3.15 | 2.29 | 1.33 | 12.50 | 41.845 | -81.1006 |
| GS | 409 | 899 | -25.00 | 42.54 | 11.82 | 13.29 | 3.20 | 3.03 | 0.99 | 8.84 | 41.720 | -81.3861 |
| GS | 158 | 40 | -25.46 | 40.32 | 13.84 | 12.81 | 3.15 | 4.43 | 0.87 | 8.84 | 41.720 | -81.3861 |
| GS | 115 | 17 | -25.46 | 40.32 | 13.84 | 12.81 | 3.15 | 4.43 | 0.87 | 8.84 | 41.720 | -81.3861 |
| GS | 152 | 32 | -23.80 | 44.88 | 14.18 | 14.45 | 3.11 | 2.35 | 1.33 | 12.50 | 41.845 | -81.1006 |
| GS | 55 | 2 | -22.31 | 45.90 | 11.06 | 14.09 | 3.26 | 4.49 | 1.00 | 8.00 | 41.974 | -82.8937 |
| GS | 165 | 43 | -24.09 | 44.12 | 13.40 | 14.23 | 3.10 | 2.91 | 1.35 | 12.50 | 41.845 | -81.1006 |
| GS | 59 | 2.3 | -23.08 | 45.46 | 16.18 | 14.10 | 3.22 | 3.16 | 1.37 | 8.00 | 41.780 | -82.7320 |
| GS | 440 | 1150 | -23.91 | 43.64 | 13.34 | 14.07 | 3.10 | 2.84 | 1.22 | 12.50 | 41.845 | -81.1006 |

**The protocols followed during sample collection and storage**

The samples were collected between May and October during the 2019 Lake Erie CSMI (Forage Task Group 2023). The intended sample size was at least 5 but with a goal of 10 individuals per basin. Fish were euthanized according to the standard operating procedure of the government agency conducting the survey, which complied with local government animal care procedures and were placed on ice for transportation to the Great Lakes Institute for Environmental Research. Upon dissection, total length and wet weight of each fish was measured and a 1–2 g skinless muscle sample was taken from below the dorsal fin and above the lateral line for stable isotope analysis. The samples were stored under -20 °C prior to the analysis.

**The stable isotope analysis**

In order to eliminate the bias in $\delta^{13}C$ in fish muscle samples compared to pure protein first the samples were lipid extracted with a 2:1 chloroform: methanol mixture. Aliquots of fish muscle were weighed to masses of 400–600 µg in tin cups for $\delta^{13}C$ and $\delta^{15}N$ analysis. Isotopic compositions of $\delta^{13}C$ and $\delta^{15}N$ were determined using a Delta V Advantage isotope ratio mass spectrometer (Thermo Scientific, Bremen, Germany) operating in continuous flow mode and coupled to an EA IsoLink IRMS (Thermo Fisher Scientific, Waltham, MA, USA). Instrument accuracy, assessed using National Institute of Standards and Technology (NIST) reference materials analyzed alongside samples, was within 0.1‰ of certified values for $\delta^{15}N$ (NIST 8573, 8547, and 8574) and $\delta^{13}C$ (NIST 8573, 8542, and 8574). Analytical precision, based on repeated measurements of four laboratory standards (NIST 1577c, tilapia muscle, USGS 40, and urea; n = 187 each) analyzed every 12 samples, was ±0.2‰ for $d_{15}N$ and ±0.1‰ for $d_{13}C$. Sample reproducibility, determined from triplicate analyses of tissue replicates (n = 71) measured every 12 samples, was ± 0.2‰ for both $\delta^{13}C$ and $\delta^{15}N$.

**Sample extraction and cleanup procedure**

A new extraction and purification procedure was developed and validated for simultaneous analysis of polycyclic aromatic nitrogen heterocycles (PANHs), polycyclic aromatic sulfur heterocycles (PASHs) and polycyclic aromatic oxygen heterocycles (PAOHs). Approximately 5 g of wet fish muscle tissue samples were freeze dried and then mixed with 3.5 g of baked DE dispersant and added to an accelerated solvent extraction (ASE) cell fitted with a fiber glass filter at the bottom of the cell. Each cell was spiked then with 10.0 µL of 1 ng/µL mixture of RIS and filled with baked Ottawa sand. Average RIS recoveries are summarised in Table SI3. Samples were extracted using 50:50 mixture of DCM: ethyl acetate using a Dionex Accelerated Solvent Extractor (ASE) 350 (ThermoFisher, Mississauga, ON). The extraction was carried out with system pressure of 1500 psi, oven temperature of 120 °C, static time of 15 min with 2 static cycles, flush volume of 60% and a nitrogen purge for 80 s at 150 psi. Extracts were dried with sodium sulfate, reduced to 5 mL and 0.5 mL was aliquoted for gravimetric lipid weight determination (Table SI5). The remaining sample volume was applied to the head of a gel permeation chromatography column (J2-scientific AccuPrepMPSTM, Columbia, Missouri, USA) packed with 60 g of S-X3 Biobeads using 50:50 (v/v) DCM: hexane as the mobile phase with the pressure of 8 psi and flow rate of 5 mL/min. The first 120 mL containing lipids and other matrices were discarded and the fraction between 120 mL to 250 mL was collected for further processing. Then the sample volume was reduced to 2 mL using a rotor evaporator. Then a mixture of 4 g of silica gel and 1 g of 5 % deactivated alumina and 0.5 g of sodium sulfate was prewashed with 10 mL of 70:30 (v/v) DCM: hexane and then 10 mL of ethyl acetate and transferred to the flat bottom flask containing the sample with 2 x 10 mL of 70:30 (v/v) DCM: hexane mixture. Samples were swirled intermittently for 30 minutes and then transferred to 60 mL test tubes via glass wool with 2 x 5 mL of 70:30 (v/v)

DCM: hexane rinse and 10 mL of ethyl acetate rinse. Sample volume was reduced to 100 µL using a gentle stream of ultra-high purity (UHP) nitrogen, spiked with 10.0 µL of 1 ng/µL of IPIS (anthracene-$d_{10}$) prior to the GC-MS/MS injection.

Samples were processed in 8 batches and in each batch two method blanks, a spiked and non spiked store-bought salmon muscle tissue samples were processed as quality control measures. Small concentrations of PACs were measured in our procedural blanks (for heterocyclic compounds 0.01 to 9.2 pg/ µL and for homocyclics 0.3 to 149.6 pg/µL except for naphthalenes which was < 500 pg/µL) and used to correct concentrations in our samples. The recovery percentages were > 60% for all analytes. Negligible amounts of analytes were detected in solvent blanks which were run between every 25 samples in the GC-MS/MS. The RIS recoveries were calculated for each sample to ensure the acceptable recovery (> 63 ± 18 %).

**Method validation protocols**

Due to the unavailability of commercial standard reference materials store bought wild salmon sample was homogenised, freeze dried and used as the matrix. three concentration levels (25, 125 and 500 pg/µL) were used for the method validation. Six matrix blanks were processed without spiking the HPACs as a quality control measure.

To check accuracy the percent recoveries were calculated by extracting 10 replicants of each concentration level spiked samples. The relative standard deviation (RSD) was calculated for the percent recovery values of each concentration level to assess the repeatability of the method. The intermediate precision was measured using three sets of 125 pg/ µL fortified samples (n=4) that were extracted on three separate days. The RSD were calculated for each sample set and single-factor ANOVA test was used to confirm the intermediate precision.

The detection limits, standard deviations (SD) were calculated based on the 10 pg/ µL spike recoveries (n=10) and blank concentrations (n=6) following the procedure in Eurachem guidelines. Then LOD (3 x SD) and LOQ (10 x SD) values were calculated. The ruggedness of the method was determined by calculating the RSDs of 125 pg/ µL spike recoveries that were extracted in two sets of spike samples. For one set (n=4) the mass of each dispersant (silica gel and 5% deactivated alumina) was deliberately increased by 0.5 g and, the second set was extracted at 140 °C instead of 120 °C (n=4).

**Table IV 4:** Performance characteristics of the extraction and purification method used for HPACs

| The Chemical Name | Selected ISTD | % Recovery | | | Repeatability (RSTD%) | | | Intermediate precision at 125 pg/ul (RSTD%) | | LOD (ng/g) | LOQ (ng/g) | Ruggedness % Recovery | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 500 pg/ul | 125 pg/ul | 25 pg/ul | 500 pg/ul | 125 pg/ul | 25 pg/ul | Day 2 | Day 3 | | | Dispersent mass | ASE Tem |
| **HPACs** | | | | | | | | | | | | | |
| Quinoline | Naphthalene d8 | 86% | 102% | 40% | 19.6 | 8.3 | 18.2 | 18.5 | 13.1 | 0.8 | 2.7 | 91% | 94% |
| Isoquinoline | Naphthalene d8 | 82% | 115% | 48% | 27.0 | 10.8 | 27.2 | 23.8 | 18.0 | 1.1 | 3.8 | 74% | 84% |
| Indole | Naphthalene d8 | 90% | 101% | 3% | 14.1 | 17.0 | 33.2 | 23.6 | 14.7 | 2.2 | 7.4 | 144% | 146% |
| 2-Methylquinoline | Naphthalene d8 | 84% | 113% | 56% | 27.2 | 11.1 | 23.8 | 21.8 | 32.8 | 1.1 | 3.6 | 85% | 87% |
| 4-Methylbenzothiophene | Fluorene d10 | 97% | 88% | 95% | 5.7 | 11.5 | 9.1 | 15.7 | 15.7 | 0.8 | 2.8 | 98% | 109% |
| 3-Methylisoquinoline | Naphthalene d8 | 81% | 106% | 64% | 35.2 | 14 | 26.8 | 22.0 | 31.5 | 1.2 | 4.1 | 68% | 77% |
| 2,3dimethylbenzothiophene | Fluorene d10 | 93% | 80% | 85% | 9.5 | 8.6 | 11.4 | 4.9 | 8.5 | 0.6 | 1.9 | 94% | 97% |
| 4-phenylpiridine | Naphthalene d8 | 83% | 121% | 55% | 32.1 | 14.6 | 27.5 | 27.9 | 22.1 | 1.5 | 5.1 | 70% | 87% |
| Dibenzofuran | Fluorene d10 | 107% | 104% | 94% | 5.7 | 3.9 | 5.9 | 3.6 | 16.0 | 0.4 | 1.3 | 137% | 133% |
| 2,3,4 trimethylbenzothiophene | Fluorene d10 | 106% | 97% | 103% | 4.2 | 6.9 | 5.6 | 6.7 | 6.1 | 0.6 | 1.9 | 127% | 125% |
| 2-Methyldibenzofuran | Phenanthrene d 10 | 123% | 121% | 123% | 7.7 | 12.6 | 6.9 | 2.6 | 8.4 | 1.3 | 4.4 | 131% | 128% |
| Dibenzothiophene | Phenanthrene d 10 | 124% | 118% | 128% | 5.6 | 4.8 | 7.0 | 2.9 | 8.0 | 0.5 | 1.7 | 142% | 136% |
| Benzo[h]quinoline | Phenanthrene d 10 | 110% | 106% | 115% | 6.5 | 6.3 | 17.2 | 4.9 | 10.1 | 0.7 | 2.2 | 113% | 129% |
| Acridine | Phenanthrene d 10 | 54% | 68% | 91% | 13.7 | 11.6 | 29.6 | 27.6 | 17.4 | 0.8 | 2.8 | 68% | 87% |
| 9-methylcarbazole | Phenanthrene d 10 | 122% | 120% | 125% | 5.7 | 5.8 | 4.1 | 7.6 | 8.8 | 0.6 | 1.9 | 130% | 137% |
| Carbazole | Phenanthrene d 10 | 130% | 132% | 122% | 6.9 | 4.3 | 9.0 | 5.4 | 20.4 | 0.5 | 1.7 | 117% | 139% |
| 4 methyl dibenzothiophene | Phenanthrene d 10 | 121% | 120% | 138% | 7.9 | 7.7 | 21.4 | 15.3 | 16.6 | 1.0 | 3.4 | 159% | 155% |
| 3-Methylbenzo[f]quinoline | Phenanthrene d 10 | 74% | 92% | 64% | 28.1 | 9.2 | 27.2 | 22.2 | 16.6 | 0.7 | 2.4 | 68% | 72% |
| 2-Methylacridine | Phenanthrene d 10 | 54% | 72% | 60% | 19.7 | 15.4 | 37.9 | 10.1 | 26.3 | 1.4 | 4.6 | 96% | 104% |
| 2,4 dimethyldibenzothiophene | Fluoranthene d11 | 133% | 109% | 107% | 7.0 | 17.7 | 22.2 | 1.4 | 9.9 | 1.8 | 6.1 | 161% | 175% |
| 3,6-Dimethylcarbazole | Fluoranthene d11 | 120% | 115% | 104% | 5.7 | 4.5 | 11.0 | 10.6 | 11.5 | 0.4 | 1.5 | 128% | 148% |
| 2,4,7 trimethyldibenzothiophene | Pyrene d11 | 110% | 105% | 79% | 10.7 | 13.7 | 32.4 | 8.2 | 2.9 | 1.3 | 4.5 | 154% | 157% |
| Benzo[b]naphtho[2,3-d]furan | Pyrene d11 | 132% | 124% | 102% | 15.4 | 12.5 | 39.8 | 5.1 | 3.2 | 1.4 | 4.8 | 151% | 137% |
| 3-tert-butyl-9H-carbazole | Pyrene d11 | 119% | 115% | 114% | 10.6 | 7.7 | 6.7 | 3.2 | 8.5 | 0.7 | 2.5 | 131% | 156% |
| 9-Phenylcarbazole | Benz(a)anthrazene d12 | 117% | 111% | 116% | 19.8 | 12.1 | 6.8 | 1.9 | 6.3 | 1.1 | 3.7 | 151% | 137% |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | Benz(a)anthrazene d12 | 100% | 88% | 113% | 7.0 | 11.7 | 13.3 | 7.1 | 22.1 | 0.9 | 2.9 | 138% | 130% |
| Benz[c]acridine | Benz(a)anthrazene d12 | 111% | 110% | 123% | 11.2 | 7.7 | 20.1 | 14.9 | 11.9 | 0.7 | 2.4 | 136% | 167% |
| 11H-Benzo[a]carbazole | Benz(a)anthrazene d12 | 116% | 112% | 95% | 12.8 | 10.1 | 18.1 | 21.8 | 7.1 | 1.0 | 3.2 | 113% | 166% |
| 7H-Benzo[c]carbazole | Benz(a)anthrazene d12 | 109% | 109% | 83% | 8.5 | 7.4 | 15.9 | 17.1 | 3.3 | 0.7 | 2.3 | 107% | 150% |
| 6-Methylbenzo[b]naphtho[2,1-d]thiophene | Benz(a)anthrazene d12 | 117% | 103% | 113% | 8.1 | 11.7 | 8.4 | 15.2 | 43.0 | 1.0 | 3.3 | 146% | 155% |
| Dinaphtho[1,2-b;1',2'-d]furan | Benzo(ghi)perylene d12 | 116% | 116% | 115% | 9.3 | 12.2 | 8.1 | 6.7 | 3.6 | 1.2 | 3.9 | 124% | 126% |
| Dibenz[c,h]acridine | Benzo(ghi)perylene-d12 | 126% | 120% | 120% | 9.4 | 10.2 | 6.3 | 9.0 | 3.2 | 1.0 | 3.3 | 130% | 136% |
| 13(H)-Dibenzo[a,i]carbazole | Benzo(ghi)perylene-d12 | 123% | 114% | 93% | 8.7 | 8.9 | 12.5 | 13.7 | 7.2 | 0.8 | 2.8 | 109% | 139% |
| **Halo-HPACs** | | | | | | | | | | | | | |
| 2,3-Dibromothiophene | Fluorene d10 | 67% | 87% | 82% | 4.7 | 5.7 | 7.8 | 7.7 | 6.8 | 0.6 | 2.1 | 56% | 54% |
| 5-Chlorobenzothiophene | Fluorene d10 | 80% | 83% | 90% | 2.7 | 7.2 | 3.5 | 6.5 | 3.9 | 0.6 | 2.1 | 78% | 73% |
| 2-chloroquinoline | Fluorene d10 | 91% | 90% | 99% | 2.4 | 6.3 | 2.6 | 5.6 | 3.8 | 0.6 | 1.9 | 89% | 87% |
| 2-Bromobenzothiophene | Fluorene d10 | 75% | 90% | 91% | 3.1 | 4.8 | 4.3 | 7.0 | 2.2 | 0.7 | 2.3 | 79% | 76% |
| 2,3-Dichlorobenzothiophene | Fluorene d10 | 77% | 85% | 90% | 3.5 | 4.5 | 3.9 | 5.2 | 2.2 | 0.5 | 1.6 | 78% | 75% |
| 2-bromoquinoline | Naphthelene d8 | 94% | 95% | 93% | 4.0 | 7.8 | 4.9 | 7.9 | 2.8 | 0.9 | 3.0 | 95% | 94% |
| 5-chloroindole | Fluorene d10 | 79% | 67% | 61% | 4.7 | 13.0 | 10.1 | 19.5 | 4.7 | 1.3 | 4.2 | 77% | 78% |
| 3-bromoisoquinoline | Fluorene d10 | 93% | 96% | 93% | 2.5 | 7.6 | 3.5 | 6.8 | 3.7 | 0.8 | 2.6 | 100% | 101% |
| 5-bromoindol | Naphthalene d8 | 94% | 84% | 67% | 6.0 | 7.3 | 5.9 | 19.4 | 5.2 | 1.4 | 4.7 | 83% | 85% |
| 2,3-Dibromobenzothiophene | Phenanthrene d10 | 74% | 80% | 86% | 4.1 | 7.4 | 4.6 | 6.7 | 6.1 | 0.6 | 2.1 | 72% | 67% |
| 9-chloroacridine | Phenanthrene d10 | 43% | 35% | 48% | 19.2 | 14.1 | 9.1 | 24.6 | 25.3 | 1.7 | 5.6 | 91% | 80% |
| 2-Bromodibenzothiophene | Phenanthrene d10 | 75% | 93% | 100% | 8.2 | 10.4 | 5.7 | 11.2 | 1.3 | 1.2 | 3.9 | 83% | 87% |
| 9-bromoacridine | Phenanthrene d10 | 42% | 37% | 47% | 20.7 | 28.0 | 12.4 | 22.8 | 20.5 | 1.2 | 4.0 | 74% | 74% |
| 3,6-dichloro-9h-carbazole | Pyrene-d11 | 99% | 99% | 94% | 3.1 | 8.1 | 6.6 | 10.5 | 5.8 | 1.1 | 3.8 | 113% | 113% |
| 2,8-Dibromodibenzothiophene | Pyrene-d11 | 99% | 87% | 97% | 3.5 | 6.8 | 4.1 | 7.2 | 4.7 | 0.7 | 2.3 | 101% | 102% |
| arbazole | Chrysene d12 | 86% | 78% | 83% | 4.6 | 9.7 | 5.1 | 9.5 | 7.7 | 0.8 | 2.7 | 96% | 97% |

**Table IV 5:** Average lipid percentage of each species.

| **Species Name** | **Average lipid content** | **Standard deviation** |
|---|---|---|
| Freshwater drum | 7.1 % | 5.6 |
| Gizzard shad | 5.7 % | 6.7 |
| Yellow perch | 0.6 % | 0.5 |
| Round goby | 4.7 % | 2.9 |
| White sucker | 2.4 % | 1.0 |
| Spottai1 shiner | 4.7 % | 2.0 |
| Emerald shiner | 1.3 % | 0.4 |
| White bass | 0.4 % | 0.3 |
| Walleye | 0.7 % | 1.0 |

**Table IV 6:** The lipid normalized concentrations (ng/ g) of each PAC groups in Lake Eire western basin samples expressed as median and mean ± standard deviation

| Species | **∑31PANHs** | **∑20PASHs** | **∑4PAOHs** | **∑16PAHs** | **∑33APAHs** | **∑104PACs** |
|---|---|---|---|---|---|---|
| Spottail Shiner | 1201 | 362 | 1329 | 111 | 3110 | 6016 |
| | 1872.9 ± 2090.7 | 600.7 ± 615.7 | 3403.2 ± 5592.3 | 313.3 ± 335.4 | 5970.9 ± 8447.2 | 12161.0 ± 16739.2 |
| Emerald Shiner | 8618 | 147 | 10 | 222 | 3976 | 16306 |
| | 12831.5 ± 11164.6 | 536.4 ± 712.7 | 52.9 ± 109.7 | 1723.9 ± 3135.0 | 4684.4 ± 3113.8 | 19829.0 ± 14295.3 |
| Yellow Perch | 10826 | 159 | 71 | 551 | 1991 | 14872 |
| | 27676.7 ± 29187.2 | 211.4 ± 155.8 | 172.4 ± 264.7 | 760.7 ± 467.0 | 2082.4 ± 1164.9 | 30903.7 ± 29369.4 |
| Gizzard Shad | 4134 | 124 | 21 | 369 | 1407 | 5780 |
| | 16954.3 ± 26629.6 | 241.0 ± 297.3 | 46.0 ± 64.4 | 656.0 ± 526.5 | 3109.6 ± 4249.3 | 21006.9 ± 30898.5 |
| White Bass | 11050 | 137 | 13 | 591 | 2837 | 14311 |
| | 13767.6 ± 9831.6 | 146.7 ± 74.5 | 18.6 ± 19.4 | 893.3 ± 1011.5 | 3748.8 ± 2648.8 | 18575.0 ± 12224.5 |
| Walleye | 2275 | 139 | 11 | 1197 | 2540 | 6982 |
| | 3297.6 ± 3602.1 | 184.7 ± 145.4 | 66.3 ± 88.8 | 1219.2 ± 662.2 | 2642.5 ± 1384.2 | 7410.3 ± 4034.0 |
| Zebra Mussels | 98128 | 8687 | 356 | 23082 | 85016 | 287780 |
| | 101104.6 ± 58501.9 | 13601.3 ± 10119.4 | 680.8 ± 697.7 | 29465.3 ± 18612.7 | 105655.7 ± 60776.6 | 250507.7 ± 128682.0 |
| Round Goby | 16908 | 778 | 88 | 2954 | 15116 | 34525 |
| | 16033.9 ± 2790.5 | 770.4 ± 339.3 | 118.4 ± 102.8 | 4034.6 ± 3791.4 | 14512.1 ± 3655.4 | 35469.3 ± 8971.1 |
| Freshwater Drum | 9345 | 4557 | 114 | 2566 | 39161 | 110301 |
| | 36365.8 ± 59546.0 | 5158.8 ± 5125.7 | 240.7 ± 298.3 | 2655.3 ± 2453.8 | 57827.5 ± 56217.5 | 102248.1 ± 70831.6 |
| White Sucker | 1938 | 92 | 100 | 1576 | 4142 | 8023 |
| | 2063.9 ± 656.1 | 183.7 ± 194.2 | 200.7 ± 306.5 | 1451.9 ± 853.5 | 4005.1 ± 2081.9 | 7905.3 ± 3336.3 |

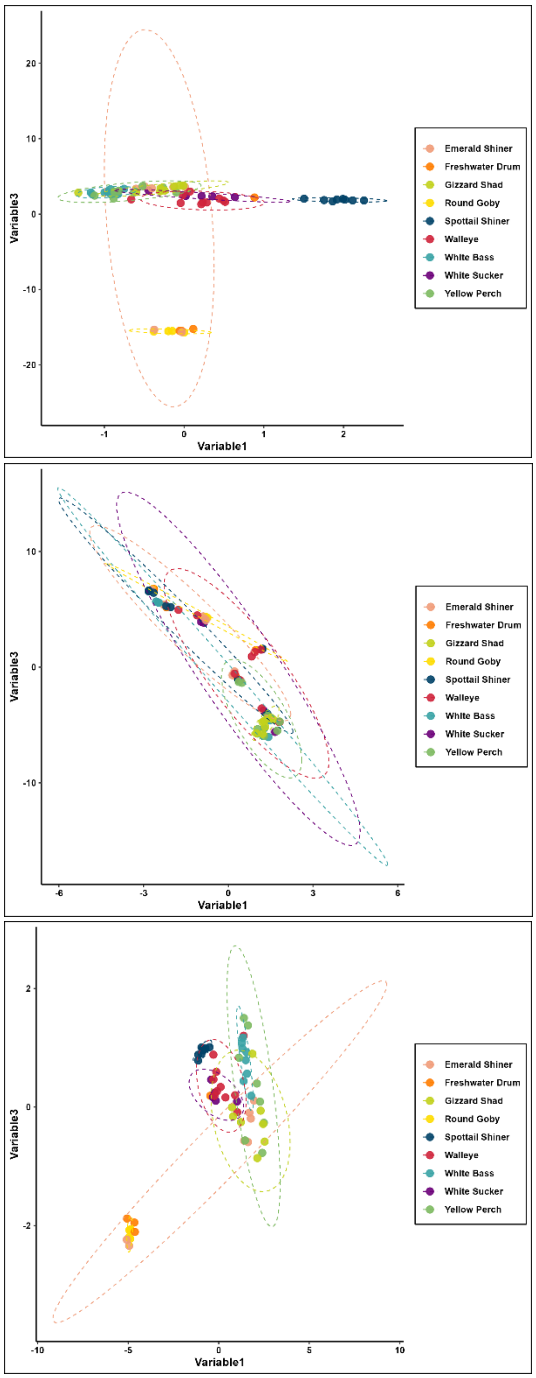



**Figure IV 1:** The UMAP analyses output graphs generated considering variable 1 and variable 3 with (a) only HPACs, (b) with PAHs and APAHs and (c) all PACs

**Table IV 7:** The calculated biomagnification factors considering major prey-predator relationships

| PACs | WB vs RG | WB vs FD | WB vs WS | WB vs SS | WB vs ES | WB vs YP | WB vs GS | WE vs RG | WE vs FD | WE vs WS | WE vs SS | WE vs ES | WE vs WB | WE vs YP | WE vs GS | YP vs RG | YP vs WS | YP vs SS | YP vs ES | YP vs GS | FD vs RG | FD vs WS | FD vs SS | FD vs ES | FD vs GS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PANH1 | 4 | 2 | 2 | 29 | 2 | 1 | 1 | 1 | 0 | 0 | 7 | 0 | 0 | 0 | 0 | 7 | 3 | 49 | 3 | 1 | 3 | 1 | 18 | 1 | 0 |
| PANH2 | 1 | 1 | 0 | 35 | 1 | 2 | 2 | 0 | 0 | 0 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 18 | 0 | 1 | 1 | 1 | 44 | 1 | 3 |
| PANH3 | 14 | 11 | 7 | 7 | 3 | 1 | 1 | 3 | 2 | 1 | 1 | 1 | 0 | 0 | 0 | 22 | 11 | 11 | 6 | 2 | 1 | 1 | 1 | 0 | 0 |
| PANH4 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 |
| PANH5 | 1451 | 135 | 30 | 1 | 2 | 3 | 1 | 325 | 30 | 7 | 0 | 0 | 0 | 1 | 0 | 573 | 12 | 0 | 1 | 1 | 11 | 0 | 0 | 0 | 0 |
| PANH6 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 3 | 0 | 0 | 6 | 22 | 255 | 20 | 2 | 1 | 3 | 32 | 3 | 0 |
| PANH7 | 84 | 10 | 8 | 19 | 1 | 0 | 0 | 12 | 1 | 1 | 3 | 0 | 0 | 0 | 0 | 412 | 40 | 92 | 5 | 1 | 9 | 1 | 2 | 0 | 0 |
| PANH8 | 0 | 0 | 0 | 1 | 1 | 1 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 3 | 2 | 3 | 6 | 8 | 31 |
| PANH9 | 517 | 1 | 0 | 533 | 3 | 1 | 2 | 19 | 0 | 0 | 19 | 0 | 0 | 0 | 0 | 504 | 0 | 518 | 3 | 2 | 486 | 0 | 500 | 3 | 2 |
| PANH10 | 3 | 2 | 1 | 21 | 2 | 0 | 1 | 7 | 5 | 3 | 51 | 5 | 2 | 1 | 3 | 8 | 3 | 62 | 6 | 3 | 1 | 1 | 10 | 1 | 1 |
| PANH11 | 270 | 0 | 5 | 0 | 0 | 0 | 0 | 3586 | 2 | 69 | 0 | 1 | 13 | 0 | 0 | 26435 | 508 | 2 | 10 | 1 | 1592 | 31 | 0 | 1 | 0 |
| PANH12 | 136 | 1 | 3 | 140 | 2 | 1 | 1 | 25 | 0 | 1 | 26 | 0 | 0 | 0 | 0 | 147 | 3 | 151 | 2 | 1 | 144 | 3 | 148 | 2 | 1 |
| PANH13 | 0 | 1 | 1 | 5 | 2 | 1 | 1 | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 1 | 1 | 1 | 6 | 2 | 2 | 1 | 2 | 9 | 3 | 3 |
| PANH14 | 200 | 0 | 2 | 1 | 0 | 0 | 2 | 168 | 0 | 1 | 1 | 0 | 1 | 0 | 2 | 406 | 3 | 3 | 1 | 5 | 641 | 5 | 4 | 1 | 7 |
| PANH15 | 1 | 2 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 0 |
| PANH16 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 3 | 2 | 0 | 1 | 3 | 17 |
| PANH17 | 39 | 0 | 0 | 0 | 0 | 0 | 0 | 143 | 0 | 1 | 0 | 1 | 4 | 1 | 1 | 249 | 1 | 0 | 2 | 2 | 947 | 5 | 2 | 7 | 9 |
| PANH18 | 2 | 0 | 2 | 0 | 0 | 0 | 0 | 3 | 0 | 3 | 0 | 0 | 1 | 0 | 0 | 39 | 36 | 1 | 1 | 1 | 76 | 72 | 3 | 1 | 3 |
| PANH19 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 3 | 0 | 0 | 0 | 9 | 0 | 2 | 2 | 0 | 28 | 1 | 6 | 5 |
| PANH20 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 3 | 8 | 4 | 2 | 0 | 2 | 5 | 3 | 1 |
| PANH21 | 0 | 0 | 0 | 0 | 0 | 4 | 1 | 2 | 2 | 1 | 2 | 1 | 4 | 15 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| PANH22 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 2 | 0 | 1 | 1 | 1 | 6 | 0 | 3 | 2 |

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PANH23 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 7 | 0 | 0 | 0 | 83 | 0 | 0 | 4 | 8 | 27693 | 50 | 41 | 1242 |
| PANH24 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 2 | 0 | 0 | 1 | 63 | 1 | 1 | 2 | 5 | 322 | 7 | 4 | 11 |
| PANH25 | 0 | 0 | 955 | 3 | 0 | 0 | 1 | 0 | 0 | 1609 | 5 | 0 | 2 | 0 | 1 | 0 | 6084 | 20 | 0 | 5 | 1 | 37504 | 121 | 2 | 29 |
| PANH26 | 0 | 0 | 8 | 72 | 0 | 1 | 1 | 0 | 0 | 2 | 15 | 0 | 0 | 0 | 0 | 1 | 13 | 121 | 0 | 1 | 8 | 181 | 1705 | 3 | 20 |
| PANH27 | 0 | 0 | 100 | 30 | 0 | 1 | 3 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 85 | 25 | 0 | 3 | 3 | 3896 | 1156 | 7 | 116 |
| PANH28 | 0 | 0 | 12 | 0 | 0 | 0 | 1 | 0 | 0 | 6 | 0 | 0 | 1 | 0 | 0 | 0 | 27 | 0 | 0 | 1 | 6 | 3608 | 45 | 15 | 169 |
| PANH29 | 17 | 0 | Inf | 0 | 2 | 0 | 0 | 235 | 0 | Inf | 0 | 24 | 14 | 0 | 1 | 490 | Inf | 0 | 50 | 2 | 746000 | Inf | 26 | 7580 | 323 |
| PANH30 | 0 | 0 | 9 | 1 | 1 | 1 | 0 | 0 | 0 | 7 | 1 | 1 | 1 | 1 | 0 | 0 | 11 | 1 | 1 | 1 | 3 | 357 | 32 | 35 | 18 |
| PANH31 | 0 | 0 | 17 | 19 | 1 | 10 | Inf | 0 | 0 | 9 | 10 | 0 | 1 | 5 | Inf | 0 | 2 | 2 | 0 | Inf | 1 | 351 | 384 | 12 | Inf |
| PASH1 | 0 | 3 | 2 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 4 | 1 | 2 | 2 | 0 | 1 | 0 | 0 | 0 |
| PASH2 | 3 | 3 | 2 | 0 | 59 | 3 | 3 | 2 | 3 | 1 | 0 | 47 | 1 | 2 | 2 | 1 | 1 | 0 | 23 | 1 | 1 | 1 | 0 | 17 | 1 |
| PASH3 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 |
| PASH4 | 182 | 0 | 0 | 4 | 1 | 0 | 1 | 97 | 0 | 0 | 2 | 1 | 1 | 0 | 0 | 381 | 1 | 8 | 3 | 2 | 408 | 1 | 8 | 3 | 2 |
| PASH5 | 0 | 0 | 9 | 2 | 0 | 0 | 0 | 0 | 0 | 17 | 3 | 1 | 2 | 1 | 1 | 1 | 25 | 5 | 1 | 1 | 1 | 39 | 7 | 2 | 2 |
| PASH6 | 0 | 0 | 2 | 0 | 0 | 1 | 1 | 0 | 1 | 3 | 0 | 1 | 1 | 1 | 1 | 0 | 4 | 1 | 1 | 1 | 1 | 5 | 1 | 1 | 2 |
| PASH7 | 3 | 0 | 35 | 0 | 1 | 0 | 36 | 1 | 0 | 13 | 0 | 0 | 0 | 0 | 13 | 15 | 157 | 2 | 4 | 161 | 17 | 180 | 2 | 5 | 186 |
| PASH8 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 116 | 2 | 138 | 148 | 0 | 0 | 0 | 589 | 9 | 699 | 1 | 1 | 1697 | 25 | 2015 | 4 |
| PASH9 | 0 | 0 | 0 | 6 | 3 | 0 | 1 | 0 | 0 | 0 | 3 | 2 | 1 | 0 | 0 | 1 | 1 | 13 | 8 | 2 | 2 | 2 | 33 | 19 | 4 |
| PASH10 | 0 | 0 | 1 | 4 | 1 | 0 | 0 | 0 | 0 | 1 | 4 | 1 | 1 | 0 | 0 | 0 | 1 | 9 | 3 | 1 | 9 | 31 | 198 | 65 | 21 |
| PASH11 | 0 | 0 | 1486 | 2 | 1 | 0 | 0 | 1 | 0 | 3233 | 4 | 2 | 2 | 1 | 0 | 1 | 3638 | 4 | 2 | 1 | 21 | 97409 | 108 | 49 | 14 |
| PASH12 | 1580 | 0 | 3 | 1 | 1 | 1 | 0 | 5614 | 0 | 9 | 2 | 2 | 4 | 2 | 1 | 2669 | 4 | 1 | 1 | 0 | 119498 | 198 | 49 | 41 | 21 |
| PASH13 | 68 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2480 | 0 | 1 | 0 | 1 | 136203 | 14 | 29 | 4 | 33 |
| PASH14 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 12276 | 13427 | 42 | 12660 |
| PASH15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | Inf | 0 | 0 | 0 | 13 | 1 | 0 | 1 | 1 | 13589 | 743 | 38 | 1168 |

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PASH16 | 0 | 0 | 69 | 0 | 0 | 3 | 8 | 0 | 0 | 32 | 0 | 0 | 0 | 1 | 4 | 0 | 24 | 0 | 0 | 3 | 1 | 4073 | 11 | 9 | 467 |
| PASH17 | 0 | 0 | 3 | 0 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 4 | 1 | 3014 | 150 | 15 | 1554 |
| PASH18 | 0 | 0 | 12 | 0 | 1 | 0 | 2 | 0 | 0 | 10 | 0 | 1 | 1 | 0 | 1 | 0 | 58 | 1 | 3 | 7 | 1 | 2497 | 27 | 135 | 322 |
| PASH19 | 0 | 0 | 498 | 4 | 591 | 34 | 171 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 15 | 0 | 17 | 5 | 1 | 1530 | 11 | 1817 | 526 |
| PASH20 | 0 | 0 | 2 | 3 | 0 | 0 | 0 | 0 | 0 | 8 | 8 | 0 | 3 | 0 | 2 | 0 | 20 | 22 | 0 | 4 | 0 | 3009 | 3291 | 9 | 621 |
| PAOH1 | 0 | 0 | 0 | 1117 | 0 | 0 | 0 | 0 | 0 | 0 | 55 | 0 | 0 | 0 | 0 | 2 | 1 | 14364 | 6 | 4 | 1 | 1 | 11867 | 5 | 3 |
| PAOH2 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 3 | 1 | 1 | 0 | 2 | 16 | 4 | 9 | 1 | 0 | 0 | 1 | 2 | 6 | 2 | 0 | 4 | 17 |
| PAOH3 | 0 | 2 | 0 | 152 | 1 | 0 | 0 | 0 | 3 | 1 | 218 | 1 | 1 | 0 | 1 | 1 | 2 | 678 | 2 | 2 | 0 | 0 | 71 | 0 | 0 |
| PAOH4 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 7 | 0 | 146 | 157 | 35 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 16 | 1 | 320 | 1 |
| PAH1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 74722 | 44 | 1 | 95737 | 86397 | 77057 | 0 | 1 | 0 | 0 | 1 | 1 | 123665 | 72 | 2 | 127529 |
| PAH2 | 180 | 201 | 169 | 0 | 201 | 196 | 175 | 1878 | 2099 | 1768 | 0 | 2099 | 10 | 2044 | 1823 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| PAH3 | 5098 | 2 | 7 | 4 | 7 | 5 | 1 | 1273 | 0 | 2 | 1 | 2 | 0 | 1 | 0 | 1120 | 1 | 1 | 1 | 0 | 2769 | 4 | 2 | 4 | 1 |
| PAH4 | 1 | 0 | 1 | 3 | 2 | 1 | 1 | 0 | 0 | 0 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 3 | 2 | 1 | 1 | 1 | 6 | 3 | 3 |
| PAH5 | 0 | 1 | 0 | 2 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 2 | 2 | 1 | 0 | 0 | 3 | 2 | 1 |
| PAH6 | 2184 | 1 | 5 | 90 | 2 | 1 | 1 | 691 | 0 | 2 | 28 | 1 | 0 | 0 | 0 | 2353 | 6 | 97 | 2 | 1 | 2576 | 6 | 106 | 2 | 1 |
| PAH7 | 1 | 11288 | 1 | 7 | 1 | 2 | 1 | 0 | 1164 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 3 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| PAH8 | 22219 | 24833 | 1 | 50 | 1 | 42 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 531 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| PAH9 | 0 | 0 | 1 | 4 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 4 | 19 | 9 | 1 | 0 | 2 | 11 | 6 | 0 |
| PAH10 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 2 | 5 | 7 | 1 | 1 | 0 | 2 | 3 | 8 | 1 | 0 | 6 | 7 | 22 | 3 |
| PAH11 | 0 | 0 | 17 | 0 | 0 | 0 | 0 | 0 | 0 | 226 | 0 | 0 | 13 | 0 | 0 | 4 | 17718 | 7 | 3 | 7 | 2 | 6696 | 3 | 1 | 3 |
| PAH12 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 10 | 6 | 0 | 0 | 0 | 1 | 1 | 27 | 1 | 1 | 8 | 8 | 207 | 8 |
| PAH13 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 4 | 3 | 0 | 0 | 0 | 5 | 2 | 21 | 1 | 0 | 12 | 5 | 51 | 3 |
| PAH14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 2 | 1 | 2 | 1 | 1 | 7 | 5 | 9 | 5 |
| PAH15 | 0 | 0 | 52 | 0 | 1 | 0 | 0 | 0 | 0 | 53 | 0 | 1 | 1 | 0 | 0 | 0 | 269 | 0 | 6 | 1 | 6 | 7370 | 5 | 156 | 35 |

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PAH16 | 15569 | 1 | 0 | 8 | 1 | 2 | 6 | 269 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7021 | 0 | 3 | 0 | 3 | 18444 | 1 | 9 | 1 | 7 |
| APAH1 | 1114 | 1245 | 1048 | 0 | 1245 | 1212 | 1081 | 9038 | 10101 | 8506 | 2 | 10101 | 8 | 9835 | 8772 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| APAH2 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 20871 | 23327 | 19643 | 1 | 23327 | 25168 | 22713 | 20257 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| APAH3 | 2849 | 0 | 2681 | 1 | 3184 | 3100 | 2765 | 17035 | 1 | 16033 | 4 | 19039 | 6 | 18538 | 16534 | 1 | 1 | 0 | 1 | 1 | 21504 | 20239 | 5 | 24034 | 20872 |
| APAH4 | 51264 | 0 | 7 | 0 | 3 | 55787 | 111 | 18538 | 0 | 3 | 0 | 1 | 0 | 20174 | 40 | 1 | 0 | 0 | 0 | 0 | 202319 | 28 | 1 | 12 | 437 |
| APAH5 | 1 | 0 | 2 | 1 | 1 | 1 | 2 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 3 | 4 | 2 | 3 | 4 |
| APAH6 | 33359 | 0 | 13 | 2 | 1 | 36302 | 19 | 6199 | 0 | 2 | 0 | 0 | 0 | 6745 | 4 | 1 | 0 | 0 | 0 | 0 | 95470 | 38 | 6 | 2 | 56 |
| APAH7 | 4141 | 0 | 54 | 2 | 1 | 3 | 2 | 763 | 0 | 10 | 0 | 0 | 0 | 1 | 0 | 1399 | 18 | 1 | 0 | 1 | 130259 | 1703 | 60 | 19 | 52 |
| APAH8 | 98514 | 0 | 1 | 20 | 1 | 2 | 1 | 12519 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 56577 | 0 | 11 | 1 | 0 | 437627 | 3 | 87 | 5 | 3 |
| APAH9 | 21663 | 0 | 7 | 3 | 0 | 1 | 1 | 6781 | 0 | 2 | 1 | 0 | 0 | 0 | 0 | 34689 | 11 | 6 | 1 | 1 | 264903 | 83 | 43 | 5 | 11 |
| APAH10 | 0 | 0 | 13 | 0 | 0 | 0 | 0 | 0 | 0 | 122 | 1 | 2 | 10 | 2 | 2 | 0 | 51 | 1 | 1 | 1 | 1 | 452 | 5 | 8 | 7 |
| APAH11 | 0 | 0 | 0 | 0 | 27 | 3 | 0 | 0 | 0 | 1 | 0 | 147 | 5 | 14 | 2 | 0 | 0 | 0 | 10 | 0 | 3 | 8 | 1 | 1084 | 15 |
| APAH12 | 1 | 1 | 1 | 2 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 2 | 2 | 2 | 0 |
| APAH13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 3 | 1 | 1 | 0 | 1 | 2 | 1 | 1 | 1 | 4 | 10 | 6 | 4 |
| APAH14 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 2 | 1 | 1 | 1 | 4 | 6 | 3 |
| APAH15 | 0 | 0 | 24 | 0 | 1 | 0 | 0 | 0 | 0 | 38 | 0 | 1 | 2 | 0 | 0 | 1 | 89 | 1 | 2 | 1 | 3 | 361 | 3 | 10 | 3 |
| APAH16 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 2 | 1 | 1 | 1 | 2 | 3 | 2 |
| APAH17 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 2 | 1 | 1 | 2 | 3 | 5 | 2 |
| APAH18 | 681 | 0 | 18 | 1 | 0 | 0 | 0 | 1638 | 0 | 43 | 3 | 1 | 2 | 1 | 1 | 1550 | 41 | 2 | 1 | 1 | 22874 | 598 | 37 | 12 | 10 |
| APAH19 | 2176 | 0 | 2048 | 2240 | 0 | 0 | 0 | 589 | 0 | 555 | 607 | 0 | 0 | 0 | 0 | 5368 | 5053 | 5526 | 1 | 0 | 68360 | 64339 | 70370 | 9 | 3 |
| APAH20 | 669 | 0 | 630 | 1 | 0 | 0 | 0 | 1380 | 0 | 1299 | 2 | 1 | 2 | 1 | 0 | 1641 | 1545 | 3 | 1 | 0 | 45049 | 42399 | 69 | 33 | 9 |
| APAH21 | 1 | 0 | 1215 | 0 | 1 | 0 | 0 | 3 | 0 | 3040 | 1 | 3 | 3 | 0 | 0 | 7 | 6960 | 2 | 6 | 1 | 111 | 103968 | 32 | 89 | 11 |
| APAH22 | 0 | 0 | 1153 | 0 | 1 | 0 | 0 | 0 | 0 | 1670 | 0 | 2 | 1 | 1 | 1 | 0 | 2982 | 0 | 3 | 1 | 3 | 72311 | 10 | 68 | 25 |

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| APAH23 | 0 | 0 | 1932 | 1 | 1 | 0 | 0 | 0 | 0 | 13968 | 6 | 5 | 7 | 1 | 1 | 0 | 10406 | 4 | 4 | 1 | 3 | 488620 | 208 | 170 | 34 |
| APAH24 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 15218 | 2 | 12 | 19497 | 5 | 1 | 0 | 2998 | 0 | 2 | 0 | 3 | 499029 | 73 | 407 | 39 |
| APAH25 | 0 | 0 | 735 | 0 | 1 | 0 | 0 | 0 | 0 | 1553 | 0 | 1 | 2 | 0 | 0 | 0 | 4463 | 1 | 3 | 1 | 39 | 428913 | 71 | 319 | 115 |
| APAH26 | 0 | 0 | 137 | 0 | 0 | 0 | 0 | 0 | 0 | 1797 | 0 | 1 | 13 | 1 | 1 | 0 | 2115 | 0 | 1 | 1 | 123 | 1288946 | 66 | 511 | 674 |
| APAH27 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 2 | 0 | 6078 | 6648 | 7218 | 7788 | 5 | 5 | 0 | 1149 | 1257 | 1365 | 1 | 9 | 21212 | 23200 | 25189 | 18 |
| APAH28 | 0 | 0 | 0 | 0 | 0 | 2 | 4 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 3 | 0 | 0 | 0 | 0 | 2 | 1 | 3 | 2 | 4 | 37 |
| APAH29 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 556 | 523 | 572 | 621 | 1 | 1 | 1 | 1 | 1 | 0 |
| APAH30 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1749 | 1646 | 1800 | 1954 | 1 | 1 | 1 | 1 | 1 | 0 |
| APAH31 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 3 | 1 | 1 | 16 | 1 | 0 | 1 | 2 | 1 | 1 | 0 | 5 | 20 | 5 | 5 | 2 |
| APAH32 | 0 | 0 | 5 | 0 | 1 | 0 | 0 | 0 | 0 | 117 | 0 | 17 | 25 | 2 | 1 | 0 | 47 | 0 | 7 | 0 | 1 | 1401 | 1 | 208 | 8 |
| APAH33 | 0 | 0 | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 224 | 0 | 5 | 24 | 1 | 1 | 0 | 342 | 0 | 7 | 1 | 1 | 2976 | 2 | 61 | 7 |
| PANHs | 1 | 0 | 5 | 6 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 2 | 12 | 14 | 2 | 1 | 2 | 15 | 18 | 3 | 2 |
| PASHs | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 6 | 24 | 8 | 10 | 19 |
| PAOHs | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 3 | 0 | 1 | 1 | 1 | 0 | 3 | 3 | 2 | 1 | 0 | 5 | 5 |
| PAHs | 0 | 0 | 0 | 2 | 0 | 1 | 1 | 0 | 0 | 1 | 3 | 1 | 1 | 1 | 1 | 0 | 0 | 2 | 0 | 1 | 1 | 2 | 8 | 2 | 4 |
| APAHs | 0 | 0 | 1 | 1 | 1 | 2 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 4 | 12 | 9 | 12 | 16 |
| PACs | 0 | 0 | 2 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 3 | 2 | 2 | 1 | 3 | 11 | 8 | 5 | 4 |

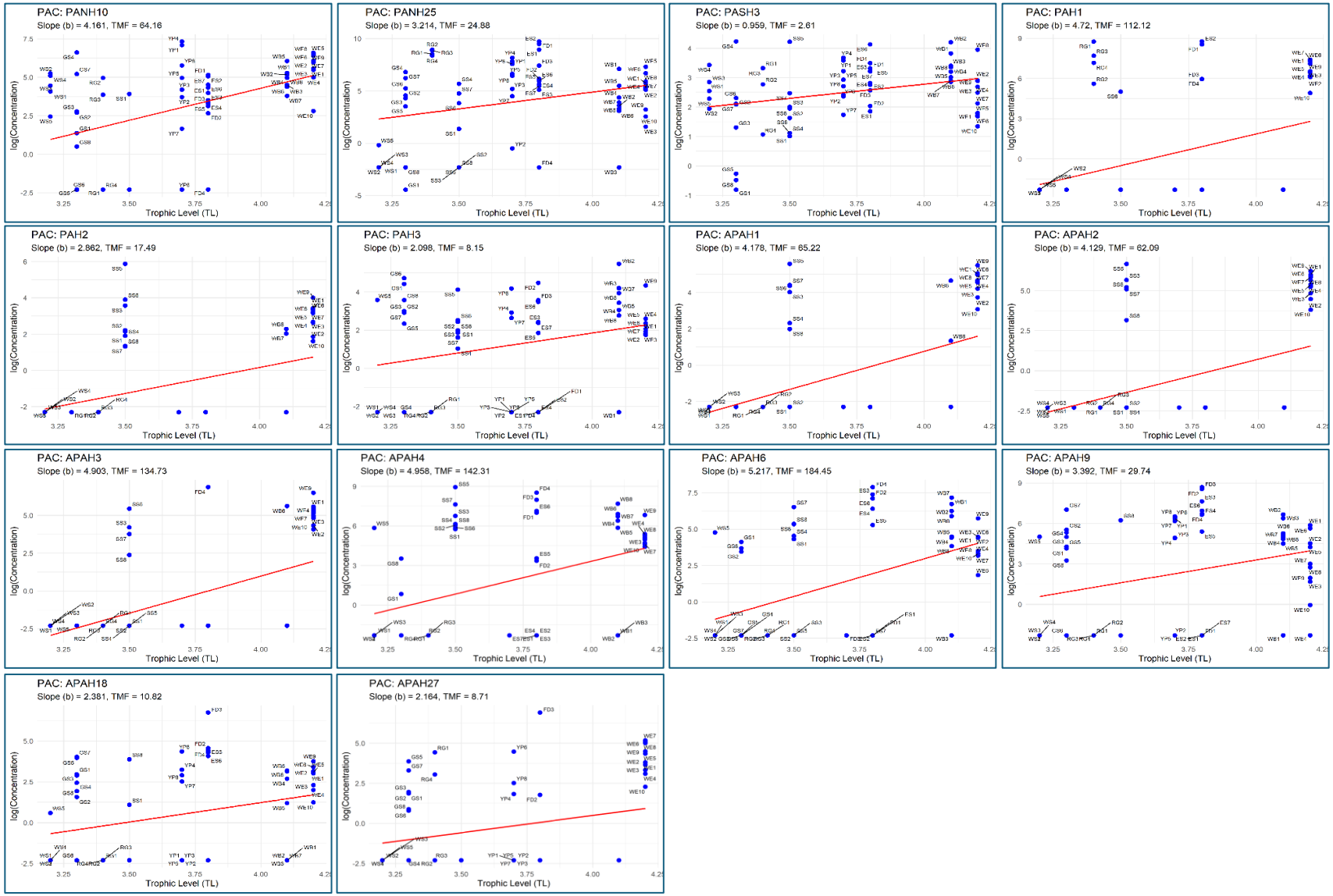



**Figure IV 2:** Plots between the $\log_{10}$concentation and tropic positions that give statistically significant biomagnifying PACs

**Table IV 8:** The calculated tropic magnification factors considering all species

| Analyte | b | TMF | p value | $R^2$ |
|---|---|---|---|---|
| Quinoline | 1.017473 | 3 | 0.18 | 0.03 |
| Isoquinoline | -0.32683 | 1 | 0.73 | 0.00 |
| Indole | 0.994473 | 3 | 0.07 | 0.06 |
| 2-Methylquinoline | -1.98401 | 0 | 0.02 | 0.09 |
| 3-Methylisoquinoline | 0.868684 | 2 | 0.31 | 0.02 |
| 2,5-Dimethyl-1-phenylpyrrole | -1.8719 | 0 | 0.03 | 0.07 |
| 4-Phenylpyridine | 1.781859 | 6 | 0.15 | 0.03 |
| 2,3-Dimethylindole | -0.23042 | 1 | 0.79 | 0.00 |
| Benzo[h]quinoline | -0.56545 | 1 | 0.46 | 0.01 |
| 1,2,3,4-Tetrahydrocarbazole | 4.161388 | 64 | 0.00 | 0.21 |
| Acridine | -0.02194 | 1 | 0.98 | 0.00 |
| 9-Methylcarbazole | 0.770358 | 2 | 0.17 | 0.03 |
| Carbazole | 0.54336 | 2 | 0.31 | 0.02 |
| 3-Methylbenzo[f]quinoline | 1.360028 | 4 | 0.05 | 0.06 |
| 2-Methylacridine | 0.58969 | 2 | 0.32 | 0.02 |
| 3,6-Dimethylcarbazole | -1.36073 | 0 | 0.13 | 0.04 |
| 3-Tert?butyl?9H-carbazole | 1.155808 | 3 | 0.11 | 0.04 |
| 9-Phenylcarbazole | -0.22212 | 1 | 0.62 | 0.00 |
| Benz[c]acridine | -1.17015 | 0 | 0.06 | 0.06 |
| 11H-Benzo[a]carbazole | -1.3515 | 0 | 0.07 | 0.05 |
| Dibenz[c,h]acridine | 0.922519 | 3 | 0.08 | 0.05 |
| 13H-Dibenzo[a,i]carbazole | -0.29729 | 1 | 0.68 | 0.00 |
| 2-Chloroquinoline | -1.27829 | 0 | 0.11 | 0.04 |
| 2-Bromoquinoline | 0.298703 | 1 | 0.74 | 0.00 |
| 5-Chloroindole | 3.213949 | 25 | 0.02 | 0.08 |
| 3-Bromoisoquinoline | 1.800891 | 6 | 0.07 | 0.05 |
| 5-Bromoindole | 1.050469 | 3 | 0.36 | 0.01 |
| 9-Chloroacridine | -0.75528 | 0 | 0.32 | 0.02 |
| 9-Bromoacridine | -0.04037 | 1 | 0.97 | 0.00 |
| 3,6-Dichlorocarbazole | 0.117711 | 1 | 0.88 | 0.00 |
| 3,6-Dibromocarbazole | 0.013999 | 1 | 0.98 | 0.00 |
| PANHs | 0.472432 | 2 | 0.36 | 0.01 |
| 4-Methylbenzothiophene | -0.76697 | 0 | 0.18 | 0.03 |
| 2,3-Dimethylbenzothiophene | -0.20976 | 1 | 0.81 | 0.00 |
| 2,3,4-Trimethylbenzothiophene | 0.958635 | 3 | 0.02 | 0.09 |
| 4-Methyldibenzothiophene | 0.073627 | 1 | 0.89 | 0.00 |
| 2,4-Dimethyldibenzothiophene | 0.761751 | 2 | 0.21 | 0.03 |
| 2,4,7-Trimethyldibenzothiophene | 0.789687 | 2 | 0.16 | 0.03 |
| 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene | 0.315917 | 1 | 0.53 | 0.01 |

| | | | | |
|---|---|---|---|---|
| 6-Methylbenzo [b]naphtho[2,1-d]thiophene | -0.45876 | 1 | 0.58 | 0.01 |
| Dibenzothiophene | -0.40363 | 1 | 0.63 | 0.00 |
| C1 Dibenzothiophene | 0.164212 | 1 | 0.87 | 0.00 |
| C2 Dibenzothiophene | 2.24642 | 9 | 0.07 | 0.05 |
| C3 Dibenzothiophene | 1.97246 | 7 | 0.11 | 0.04 |
| C4 Dibenzothiophene | -1.48251 | 0 | 0.25 | 0.02 |
| 2,3-Dibromothiophene | -0.17618 | 1 | 0.78 | 0.00 |
| 5-Chlorobenzothiophene | -1.33192 | 0 | 0.12 | 0.04 |
| 2-Bromobenzothiophene | 0.85141 | 2 | 0.26 | 0.02 |
| 2,3-Dichlorobenzothiophene | -0.11039 | 1 | 0.87 | 0.00 |
| 2,3-Dibromobenzothiophene | 0.152256 | 1 | 0.79 | 0.00 |
| 2-Bromodibenzothiophene | -0.00432 | 1 | 0.99 | 0.00 |
| 2,8-Dibromodibenzohiophene | 0.172141 | 1 | 0.71 | 0.00 |
| PASHs | -0.14713 | 1 | 0.73 | 0.00 |
| Dibenzofuran | -2.74016 | 0 | 0.02 | 0.09 |
| 2-Methyldibenzofuran | -2.12443 | 0 | 0.07 | 0.05 |
| Benzo[b]naphtho[2,3-d]furan | 0.962132 | 3 | 0.19 | 0.03 |
| Dinaphtho[1,2-b;1',2'-d]furan | -0.71915 | 0 | 0.28 | 0.02 |
| PAOHs | -1.92792 | 0 | 0.02 | 0.08 |
| Naphthalene | 4.719589 | 112 | 0.00 | 0.14 |
| Acenapthene | 2.86163 | 17 | 0.00 | 0.15 |
| Acenapthylene | 2.098483 | 8 | 0.03 | 0.08 |
| Fluorene | 0.185303 | 1 | 0.78 | 0.00 |
| Phenanthrene | -0.40132 | 1 | 0.65 | 0.00 |
| Anthracene | 0.299065 | 1 | 0.77 | 0.00 |
| Fluoranthene | -1.18399 | 0 | 0.32 | 0.02 |
| Pyrene | -1.49428 | 0 | 0.27 | 0.02 |
| Benz[a]anthracene | -2.58843 | 0 | 0.02 | 0.09 |
| Chrysene | -0.74001 | 0 | 0.47 | 0.01 |
| Benzo[a]pyrene | -0.93948 | 0 | 0.41 | 0.01 |
| Benzo[b]fluoranthene | -1.85475 | 0 | 0.07 | 0.05 |
| Benzo[k]fluoranthene | -2.05383 | 0 | 0.03 | 0.07 |
| Indeno[1,2,3-c,d]pyrene | -2.01524 | 0 | 0.04 | 0.07 |
| Dibenzo[a,h]anthracene | -1.24799 | 0 | 0.12 | 0.04 |
| Benzo[g,h,i]perylene | -1.20684 | 0 | 0.35 | 0.01 |
| PAHs | 0.300502 | 1 | 0.51 | 0.01 |
| 1-Methylnaphthalene | 4.177776 | 65 | 0.00 | 0.23 |
| 2-Methylnaphthalene | 4.128511 | 62 | 0.00 | 0.18 |
| C2 Naphthalene | 4.903259 | 135 | 0.00 | 0.26 |
| C3-Nap total | 4.957974 | 142 | 0.00 | 0.16 |
| C4-Nap total | 0.416004 | 2 | 0.35 | 0.01 |
| C1 Fluorene | 5.217385 | 184 | 0.00 | 0.21 |

| | | | | |
|---|---|---|---|---|
| C2 Fluorene | 0.40335 | 1 | 0.74 | 0.00 |
| C3 Fluorene | -0.76547 | 0 | 0.64 | 0.00 |
| C4 Fluorene | 3.392393 | 30 | 0.02 | 0.09 |
| 1,7-Dimethylphenanthrene | 1.172038 | 3 | 0.16 | 0.03 |
| 1,8-Dimethylphenanthrene | 0.026906 | 1 | 0.97 | 0.00 |
| 1-Methylphenanthrene | -0.94843 | 0 | 0.16 | 0.03 |
| 2,6-Dimethylphenanthrene | -0.55373 | 1 | 0.43 | 0.01 |
| 2-Methylphenanthrene | -1.20908 | 0 | 0.10 | 0.05 |
| 3,6-Dimethylphenanthrene | -0.00306 | 1 | 1.00 | 0.00 |
| 3-Methylphenanthrene | -1.40032 | 0 | 0.03 | 0.07 |
| 9/4-Methylphenanthrene | -1.24394 | 0 | 0.02 | 0.09 |
| C2 Phenanthrene | 2.380977 | 11 | 0.02 | 0.08 |
| C3-Phe total | 0.15324 | 1 | 0.91 | 0.00 |
| C4-Phe total | 1.523635 | 5 | 0.16 | 0.03 |
| C1 Pyrene | 1.068809 | 3 | 0.38 | 0.01 |
| C2 Pyrene | 0.590654 | 2 | 0.63 | 0.00 |
| C3 Pyrene | 0.927161 | 3 | 0.55 | 0.01 |
| C4 Pyrene | 0.121603 | 1 | 0.94 | 0.00 |
| C1 Benzo[a]pyrene | -0.48266 | 1 | 0.71 | 0.00 |
| C2 Benzo[a]pyrene | 0.4963 | 2 | 0.72 | 0.00 |
| C1 Chrysene | 2.164271 | 9 | 0.05 | 0.06 |
| C2 Chrysene | -0.31132 | 1 | 0.59 | 0.00 |
| C3-Chr total | -0.37916 | 1 | 0.37 | 0.01 |
| C4-Chr total | -0.41486 | 1 | 0.39 | 0.01 |
| Retene | -0.47931 | 1 | 0.58 | 0.01 |
| 1,2-BNT | -0.93298 | 0 | 0.25 | 0.02 |
| 2,3-BNT | -0.94608 | 0 | 0.23 | 0.02 |
| APAHs | -0.11558 | 1 | 0.79 | 0.00 |

# APPENDIX V: ADDITIONAL METRICS FROM CHAPTER 8: COMPREHENSIVE DATA FILTERING APPROACHES FOR NON-TARGETED ANALYSIS OF POLYCYCLIC AROMATIC COMPOUNDS

**Table V 1:** The targeted analytes, internal standards and their CAS numbers

| **Polyaromatic Nitrogen Heterocycles** | **Polyaromatic Sulfur and Oxygen Heterocycles** | **Halogenated Polycyclic Aromatic Hydrocarbons** | **Polycyclic Aromatic Hydrocarbons** | **Alkylated Polycyclic Aromatic Hydrocarbons** | **Internal Standards** |
|---|---|---|---|---|---|
| Quinoline - 91-22-5 | 4-Methylbenzothiophene - 14315-11-8 | 9-Chlorofluorene – 5728-80-7 | Naphthalene- 91-20-3 | 1-Methylnaphthalene – 90-12-0 | Naphthalene-d8 - 1146-65-2 |
| Isoquinoline - 119-65-3 | 2,3-Dimethylbenzothiophene - 4923-91-5 | 5-Bromoacenaphthene - 63074-49-3 | Acenapthene- 83-32-9 | 2-Methylnaphthalene – 91-57-6 | Fluorene-d10 - 81103-79-9 |
| Indole - 120-72-9 | 2,3,4-Trimethylbenzothiophene 70021-44-2 | 9-Bromofluorene - 1940-57-4 | Acenapthylene- 208-96-8 | 1,4-Dimethylnaphthalene - 571-58-4 | Phenanthrene-d10 - 1517-22-2 |
| 2-Methylquinoline – 91-63-4 | 3-Methyldibenzothiophene - 16587-52-3 | 1-Chloroanthracene - 92902-83-5 | Fluorene- 86-73-7 | 2,3,5-Trimethylnaphthalene - 2245-38-7 | Anthracene-d10 - 1719-06-8 |
| 3-Methylisoquinoline - 1125-80-0 | 2,4-Dimethyldibenzothiophene - 31317-18-7 | 2,7-Dichlorofluorene - 24463-18-3 | Phenanthrene- 1985-01-08 | 1,4,6,7-Tetramethylnaphthalene - 6682-71-9 | Fluoranthene-d10 - 93951-69-0 |
| 2,5-Dimethyl-1-phenylpyrrole - 83-24-9 | 2,4,7-Trimethyldibenzothiophene - 216983-03-8 | 1,2-Dibromoacenaphthylene - 13019-33-5 | Anthracene - 120-12-7 | 1-Methylfluorene - 1730-37-6 | Pyrene-d10 - 1718-52-1 |
| 4-Phenylpyridine – 939-23-1 | 7,8,9,10-Tetrahydrobenzo[b]naphtho[2,3-d]thiophene - 39097-15-9 | 9-Bromoanthracene - 1564-64-3 | Fluoranthene - 206-44-0 | 9-Ethylfluorene - 1821-35-8 | Chrysene-d12 - 1719-03-5 |
| 2,3-Dimethylindole – 91-55-4 | 6-Methylbenzo [b]naphtho[2,1-d]thiophene - 4567-51-5 | 1,4-Dichloroanthracene - 31310-26-0 | Pyrene - 129-00-0 | 9-Propylfluorene – 4037-45-0 | Benz[a]anthracene-d12 - 1718-53-2 |
| Benzo[h]quinoline – 230-27-3 | Dibenzothiophene - 884-12-500 | 5,6-Dibromo-1,2-dihydroacenaphthylene - 19190-91-1 | Benz[a]anthracene – 56-55-3 | 9-Butylfluorene | Benzo[b]fluoranthene-d12 - 93951-98-5 |
| 1,2,3,4-Tetrahydrocarbazole - 942-01-8 | 3-Methyldibenzothiophene - 20928-02-3 | 2,7-Dibromofluoranthene | Chrysene - 218-01-9 | 1,7-Dimethylphenanthrene - 483-87-4 | Benzo[k]fluoranthene-d12 - 93952-01-3 |
| Acridine - 260-94-6 | 2,2-Dimethyldibenzothiophene - 31317-18-7 | 3-Bromofluoranthene - 13438-50-1 | Benzo[a]pyrene - 50-32-8 | 1,8-Dimethylphenanthrene - 7372-87-4 | Benzo[g,h,i]perylene-d12 - 93952-00-2 |
| 9-Methylcarbazole - 1484-12-4 | 4-Propyldienzothiophene - 132034-86-7 | 1-Bromopyrene - 1714-29-0 | Benzo[b]fluoranthene - 205-99-2 | 1-Methylphenanthrene - 832-69-9 | |
| Carbazole - 86-74-8 | 4,6-Diethyldibenzothiophene - 132034-91-4 | 4-Bromobenzo[a]anthracene - 57332-45-7 | Benzo[k]fluoranthene - 207-08-9 | 2,6-Dimethylphenanthrene - 17980-16-4 | |

| | | | | | |
|---|---|---|---|---|---|
| 3-Methylbenzo[f]quinoline - 1985-06-03 | 2,3-Dibromothiophene – 3140-93-0 | 1,6-Dromopyrene - 75365-24-5 | Indeno[1,2,3-cd]pyrene - 193-39-5 | 2-Methylphenanthrene - 2531-84-2 | |
| 2-Methylacridine – 613-15-0 | 5-Chlorobenzothiophene - 20532-33-6 | | Dibenzo[a,h]anthracene - 53-70-3 | 3,6-Dimethylphenanthrene - 1576-67-6 | |
| 3,6-Dimethylcarbazole - 5599-50-8 | 2-Bromobenzothiophene - 5394-13-8 | | Benzo[g,h,i]perylene - 191-24-2 | 3-Methylphenanthrene - 832-71-3 | |
| 3-Tert-butyl-9H-carbazole - 22401-74-7 | 2,3-Dichlorobenzothiophene - 523-97-7 | | | 9/4-Methylphenanthrene - 883-20-5 | |
| 9-Phenylcarbazole - 1150-62-5 | 2,3-Dibromobenzothiophene - 6287-82-7 | | | 1,3-Dimethylphenantherene - 1202-70-8 | |
| Benz[c]acridine – 225-51-4 | 2-Bromodibenzothiophene - 22439-61-8 | | | 1,2,6-Trimethylphenantherene - 30436-55-6 | |
| 11H-Benzo[a]carbazole - 239-01-0 | 2,8-Dibromodibenzohiophene - 31574-87-5 | | | 1,2,6,9-Tetramethylphenantherene - 204256-39-3 | |
| Dibenz[c,h]acridine - 224-53-3 | 1,2-BNT - 205-43-6 | | | 1-Methylpyrene - 2381-21-7 | |
| 13H-Dibenzo[a,i]carbazole - 239-64-5 | 2,3- BNT- 243-46-9 | | | 1-Ethylpyrene - 1694-09-3 | |
| 2-Chloroquinoline – 612-62-4 | Dibenzofuran – 132-64-9 | | | 1-Propylpyrene – 42211-33-6 | |
| 2-Bromoquinoline - 2005-43-8 | 2-Methyldibenzofuran - 4265-25-2 | | | 1-Butylpyrene - 35980-18-8 | |
| 5-Chloroindole – 17422-32-1 | Benzo[b]naphtho[2,3-d]furan - 243-42-5 | | | 6-Methylbenzo[a]pyrene - 7496-02-8 | |
| 3-Bromoisoquinoline - 34784-02-6 | Dinaphtho[1,2-b;1',2'-d]furan - 68518-20-7 | | | 7,10-Dimethylbenzo[a]pyrene - 63104-33-6 | |
| 5-Bromoindole – 10075-50-0 | | | | 5-Methylchrysene – 1705-85-7 | |
| 9-Chloroacridine - 1207-69-8 | | | | 5-Ethylchrysene - 1206-79-7 | |
| 9-Bromoacridine – 4357-57-7 | | | | 6-Propylchrysene – 6901-41-4 | |
| 3,6-Dichlorocarbazole - 6825-20-3 | | | | 6-Butylchrysene - 6901-71-9 | |
| 3,6-Dibromocarbazole- 6825-20-3 | | | | Retene - 483-65-8 | |

**The protocols followed during sample collection and storage**

The samples were collected between May and October during the 2019 Lake Erie CSMI (Forage Task Group 2023). The intended sample size was at least 5 but with a goal of 10 individuals per basin. Fish were euthanized according to the standard operating procedure of the government agencies conducting the survey (Ontario Ministry of the Natural Resources, Ohio Department of Natural Resources and US Geological Service), which complied with local government animal care procedures and were placed on ice for transportation to the Great Lakes Institute for Environmental Research. Upon dissection, the total length and wet weight of each fish were measured, and a 1–2 g skinless muscle sample was taken from below the dorsal fin and above the lateral line for stable isotope analysis. The samples were stored under -20 °C prior to the analysis.

**Table V 2:** Sample names and sampling information for sediments used in this study (HH - Hamilton Harbour)

| Sample Name | Sampling Year | Locations | | Sample Description | Sampling Location Name |
|---|---|---|---|---|---|
| | | Latitude (° N) | Longitude (° W) | | |
| HH_Sed1 | 2014 | 43.287 | -79.796 | Bottom Sediments | Hamilton Harbour Station near Windemere |
| HH_Sed2 | 2014 | 43.279 | -79.875 | Suspended Sediments | Hamilton Harbour Station 9031 near Cootes |
| YP | 2019 | - | - | Yellow Perch Fish Muscle Tissue | Western Basin Lake Erie |
| WB | 2019 | - | - | White Bass Fish Muscle Tissue | Western Basin Lake Erie |

**Biota extraction and cleanup procedures**

A new extraction and purification procedure was developed and validated for simultaneous analysis of polycyclic aromatic nitrogen heterocycles (PANHs), polycyclic aromatic sulfur heterocycles (PASHs), and polycyclic aromatic oxygen heterocycles (PAOHs). Approximately 5 g of wet fish muscle tissue samples were dried using a LABCONCO freeze drier for 48h and then mixed with 3.5 g of baked DE dispersant and added to an accelerated solvent extraction (ASE) cell fitted with a fiber glass filter at the bottom of the cell. Each cell was spiked then with 10.0 µL of a 1 ng/µL mixture of RIS and filled with baked Ottawa sand to fill the void volume. Samples were extracted using a 50:50 mixture of DCM: ethyl acetate using a Dionex Accelerated Solvent Extractor (ASE) 350 (ThermoFisher, Mississauga, ON). The extraction was carried out with system pressure of 1500 psi, oven temperature of 120 $^{\circ}$C, static time of 15 min with 2 static cycles, flush volume of 60% and a nitrogen purge for 80 s at 150 psi. Samples were dried by directly adding sodium sulfate into sample extracts and filtering through glass wool. Then volume was reduced to 5 mL and 0.5 mL was aliquoted for gravimetric lipid weight determination (Table SI4). The remaining sample volume was applied to the head of a gel permeation chromatography (GPC) column (J2-scientific AccuPrepMPSTM, Columbia, Missouri, USA) packed with 60 g of S-X3 Biobeads using 50:50 (v/v) DCM: hexane as the mobile phase with the pressure of 8 psi and flow rate of 5 mL/min. The GPC fractions were determined by spiking native analytical standards into 1 g of cod liver oil and running through the GPC column using the above solvent system. The first 120 mL containing lipids and other matrices was discarded, and the fraction between 120 mL and 250 mL was collected for further processing. Then the sample volume was reduced to 2 mL using a rotor evaporator. Then a mixture of 4 g of silica gel and 1 g of 5 % deactivated alumina and 0.5 g of sodium sulfate was prewashed with 10 mL of 70:30 (v/v) DCM: hexane and then 10 mL of ethyl acetate and transferred to the flat bottom flask containing the sample with 2 x 10 mL of 70:30 (v/v) DCM:

hexane mixture. Samples were swirled intermittently for 30 minutes, then transferred to 60 mL test tubes via glass wool, using 2 x 5 mL of 70:30 (v/v) DCM:hexane rinse and 10 mL of ethyl acetate rinse. Sample volume was reduced to 100 µL using a gentle stream of ultra-high purity (UHP) nitrogen, spiked with 10.0 µL of 1 ng/µL of IPIS (anthracene-$d_{10}$) prior to the GC-MS/MS injection.

**Sediment extraction and cleanup procedures**

Dried sediment samples were extracted using a Dionex Accelerated Solvent Extractor 350 (ThermoFisher, Mississauga, ON). The one-step layerless extraction for PAHs and APAHs, described in Xia et al.,[31] was modified to extract all three types of HPACs. A fiber glass filter was placed in the bottom of the 34 ml ASE tubes and a mixture of dispersant was added to the cell. The dispersant mixture contained 5 g of baked silica gel, 4.5 g of 5% deactivated alumina and 1 g of reduced copper. Following this, 1.00 g of sediment sample was mixed with 1.5 g of baked diatomaceous earth (DE) dispersant and added to the ASE cell. Samples were then spiked with 10.0 µL of 10 ng/µL mixture of recovery internal standards (RIS) prior to filling the cell with baked Ottawa sand baked at 600 °C for 6 h.

Several extraction solvents, either alone or in combination, were tested in our study. Compared to dichloromethane (DCM): hexane mixtures and only DCM, a 50:50 mixture of DCM: ethyl acetate gave the most acceptable recoveries for the largest number of our targeted compounds. The ASE extraction parameters were as follows: system pressure 1500 psi, oven temperature 120 °C, static time of 15 min with 2 static cycles, 1:1 DCM: ethyl acetate mixture as the extraction solvent, 60% flush volume and a nitrogen purge for 80 s at 150 psi. Extracts were then dried by passing them through a glass funnel containing sodium sulphate and reduced in volume to 100 µL using a gentle stream of ultra-high purity (UHP) nitrogen. Finally, extracts

were transferred to a GC-vial, spiked with 10.0 µL of 10 ng/µL of instrument performance internal standard (IPIS), and diluted to 1mL final volume with hexane.

**Method validation protocols**

Due to the unavailability of commercial standard reference materials store bought wild salmon sample was homogenised, freeze dried and used as the matrix. three concentration levels (25, 125 and 500 pg/µL) were used for the method validation. Six matrix blanks were processed without spiking the HPACs as a quality control measure.

To check accuracy the percent recoveries were calculated by extracting 10 replicants of each concentration level spiked samples. The relative standard deviation (RSD) was calculated for the percent recovery values of each concentration level to assess the repeatability of the method. The intermediate precision was measured using three sets of 125 pg/ µL fortified samples (n=4) that were extracted on three separate days. The RSD were calculated for each sample set and single-factor ANOVA test was used to confirm the intermediate precision.

The detection limits, standard deviations (SD) were calculated based on the 10 pg/ µL spike recoveries (n=10) and blank concentrations (n=6) following the procedure in Eurachem guidelines. Then LOD (3 x SD) and LOQ (10 x SD) values were calculated. The ruggedness of the method was determined by calculating the RSDs of 125 pg/ µL spike recoveries that were extracted in two sets of spike samples. For one set (n=4) the mass of each dispersant (silica gel and 5% deactivated alumina) was deliberately increased by 0.5 g and, the second set was extracted at 140 °C instead of 120 °C (n=4).

**ChromaToF data processing method parameters**

The GC x GC mass calibration was carried out for all samples using the PFTBA peaks. Then, the peak finding was carried out S/N at 10 and peak quality at 0.7. The PFTBA and $CO_2$ masses were removed as background ions. Spectral similarity was done using NIST library and user created library at 3 ppm mass accuracy. The names were assigned at 750 minimum similarity scores. Formula computation was done using mass difference at 5 ppm accuracy and minimum double bond count was set up to 8 and C, H, N, S, O, Cl, Br were the allowed elements.

**Table V 3:** PACs library created using the authentic standards

| Compound | Acronyms | formula | Exact mass | 1st Dimension Time (s) | 2nd Dimension Time (s) | Base Mass |
|---|---|---|---|---|---|---|
| naphthalene-d8 | Nap-d8 | C10D8 | 136.1128 | 879 | 2.932 | 136.11 |
| naphthalene | Nap | C10H8 | 128.0626 | 884 | 0.104 | 128.06 |
| 2,3-dibromothiophene | 2,3-Br2-BT | C4H2Br2S | 239.8185 | 887 | 0.103 | 241.82 |
| benzo[b]thiophene | BbT | C8H6S | 134.0191 | 901 | 0.14 | 134.02 |
| quinoline | Qui | C9H7N | 129.0578 | 981 | 0.537 | 129.06 |
| isoquinoline | IQui | C9H7N | 129.0578 | 1029 | 0.588 | 129.06 |
| indole | Ind | C8H7N | 117.0578 | 1084 | 0.991 | 117.06 |
| 2-methylnaphthalene | 2-MeNap | C11H10 | 142.0783 | 1096 | 0.223 | 141.07 |
| 2-methylquinoline | 2-MeQui | C10H9N | 143.0735 | 1118 | 0.612 | 143.07 |
| 4-methylbenzothiophene | 4-MeBT | C9H8S | 148.0347 | 1129 | 0.516 | 147.03 |
| 1-methylnaphthalene | 1-MeNap | C11H10 | 142.0783 | 1132 | 0.386 | 141.07 |
| 3-methylisoquinoline | 3-MeIQui | C10H9N | 143.0735 | 1159 | 0.606 | 143.07 |
| 5-chlorobenzothiophene | 5-Cl-BT | C8H5ClS | 168.98 | 1298 | 0.889 | 167.98 |
| 2,3-dimethylbenzothiophene | 2,3-DiMeBT | C10H10S | 162.0504 | 1376 | 0.657 | 147.03 |
| 2-chloroquinoline | 2-Cl-Qui | C9H6ClN | 163.0187 | 1379 | 1.234 | 128.05 |
| 1,4-dimethylnaphthalene | 1,4-DiMeNap | C12H12 | 156.0939 | 1403 | 0.631 | 141.07 |
| 2-bromobenzothiophene | 2-Br-BT | C8H5BrS | 212.9294 | 1411 | 0.922 | 211.93 |
| acenaphthylene-d8 | Acy-d8 | C12D8 | 160.1128 | 1422 | 1.086 | 160.11 |
| acenaphthylene | Acy | C12H8 | 152.0626 | 1426 | 1.127 | 152.06 |
| acenaphthene-d10 | Ace-d10 | C12D10 | 164.141 | 1489 | 1.003 | 162.13 |
| acenaphthene | Ace | C12H10 | 154.0783 | 1501 | 1.04 | 153.07 |
| 2,3-dimethylindole | 2,3-DiMeInd | C10H11N | 145.0891 | 1507 | 1.28 | 144.08 |
| 2,3-dichlorobenzo[b]thiophene | 2,3-Cl2-BT | C8H4Cl2S | 202.941 | 1536 | 0.771 | 201.94 |
| dibenzofuran | DBF | C12H8O | 168.0575 | 1569 | 0.988 | 168.06 |
| 2-bromoquinoline | 2-Br-Qui | C9H6BrN | 206.9681 | 1604 | 1.697 | 128.05 |
| 5-chloroindole | 5-Cl-Ind | C8H6ClN | 151.0187 | 1624 | 1.768 | 151.02 |
| 2,3,5-trimethylnaphthalene | 2,3,5-TriMeNap | C13H14 | 170.1096 | 1672 | 0.643 | 155.09 |
| fluorene-d10 | Flu-d10 | C13D10 | 176.141 | 1704 | 1.146 | 174.13 |
| 3-bromoisoquinoline | 3-Br-IQui | C9H6BrN | 206.9681 | 1711 | 1.895 | 128.05 |
| fluorene | Flu | C13H10 | 166.0783 | 1716 | 1.171 | 165.07 |
| 2,3,4-trimethylbenzothiophene | 2,3,4-TriMeBT | C11H12S | 176.066 | 1722 | 0.963 | 161.04 |
| 2-methyldibenzofuran | 2-MeDBF | C13H10O | 182.0732 | 1825 | 0.968 | 181.06 |
| 5-bromoindole | 5-Br-Ind | C8H6BrN | 194.9681 | 1842 | 2.242 | 116.05 |
| 9-ethylfluorene | 9-EtFlu | C15H14 | 194.1096 | 1917 | 0.963 | 165.07 |
| 1,4,6,7-tetramethylnaphthalene | 1,4,6,7-TetraMeNap | C14H16 | 184.1252 | 1933 | 0.66 | 169.1 |
| 1-methylfluorene | 1-MeFlu | C14H12 | 180.0939 | 1977 | 1.203 | 165.07 |
| dibenzothiophene | DBT | C12H8S | 184.0347 | 2078 | 1.806 | 184.03 |
| 9-propylfluorene | 9-PropFlu | C16H16 | 208.1252 | 2109 | 0.877 | 165.07 |
| phenanthrene-d10 | Phe-d10 | C14D10 | 188.141 | 2127 | 1.834 | 188.14 |
| phenanthrene | Phe | C14H10 | 178.0783 | 2135 | 1.834 | 178.08 |
| anthracene-d10 | Ant-d10 | C14D10 | 188.141 | 2147 | 1.785 | 188.14 |
| benzo[h]quinoline | BhQui | C13H9N | 179.0735 | 2153 | 2.137 | 179.07 |
| anthracene | Ant | C14H10 | 178.0783 | 2154 | 1.811 | 178.08 |
| 9-chlorofluorene | 9-Cl-Flu | C13H9Cl | 200.0394 | 2172 | 1.629 | 165.07 |
| 1,2,3,4-tetrahydrocarbazole | 1,2,3,4-TetHCar | C12H13N | 171.1048 | 2176 | 2.077 | 143.07 |
| acridine | Acr | C13H9N | 179.0735 | 2175 | 2.154 | 179.07 |
| 5-bromoacenaphthene | 5-Br-Ace | C12H9Br | 231.9914 | 2201 | 1.749 | 153.07 |
| 9-methylcarbazole | 9-MeCar | C13H11N | 181.0891 | 2221 | 2.134 | 180.08 |
| carbazole | Car | C12H9N | 167.0735 | 2245 | 2.503 | 167.07 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 2-methyldibenzothiophene | 2-MeDBT | C13H10S | 198.0504 | 2282 | 1.663 | 198.05 |
| 9-butylfluorene | 9-ButFlu | C17H18 | 222.1409 | 2292 | 0.853 | 165.07 |
| 3-methyldibenzothiophene | 3-MeDBT | C13H10S | 198.0504 | 2321 | 1.629 | 198.05 |
| 3-methylphenanthrene | 3-MePhe | C15H12 | 192.0939 | 2363 | 1.654 | 192.09 |
| 3-methylbenzo[f]quinoline | 3-MeBfQui | C14H11N | 193.0891 | 2364 | 2.037 | 193.09 |
| 2-methylphenanthrene | 2-MePhe | C15H12 | 192.0939 | 2375 | 1.753 | 192.09 |
| 9-bromofluorene | 9-Br-Flu | C13H9Br | 243.9888 | 2387 | 1.809 | 165.07 |
| 9-methylphenanthrene | 9-MePhe | C15H12 | 192.0939 | 2414 | 1.843 | 192.09 |
| 2-methylacridine | 2-MeAcr | C14H11N | 193.0891 | 2418 | 1.983 | 193.09 |
| 1-methylphenanthrene | 1-MePhe | C15H12 | 192.0939 | 2427 | 1.878 | 192.09 |
| 9-chloroacridine | 9-Cl-Acr | C13H8ClN | 213.0343 | 2502 | 1.607 | 213.03 |
| 2,4-dimethyldibenzothiophene | 2,4-DiMeDBT | C14H12S | 212.066 | 2545 | 0.977 | 212.07 |
| 2,4-dimethylbenzo[h]quinoline | 2,4-DiMeBhQ | C15H13N | 207.1048 | 2556 | 1 | 207.1 |
| 1-chloroanthracene | 1-Cl-Ant | C14H9Cl | 212.0394 | 2556 | 1.091 | 212.04 |
| 3,6-dimethylphenanthrene | 3,6-DiMePhe | C16H14 | 206.1096 | 2565 | 0.693 | 206.11 |
| 2,6-dimethylphenanthrene | 2,6-DiMePhe | C16H14 | 206.1096 | 2579 | 0.663 | 206.11 |
| 2,7-dichlorofluorene | 2,7-Cl2-Flu | C13H8Cl2 | 234.0004 | 2592 | 0.808 | 199.03 |
| 1,2-dibromoacenaphthylene | 1,2-Br2-Acy | C12H6Br2 | 309.8895 | 2596 | 1 | 150.05 |
| 4-propyldibenzothiophene | 4-PropDBT | C15H14S | 226.0817 | 2606 | 0.48 | 197.04 |
| 1,3-dimethylphenanthrene | 1,3-DiMePhe | C16H14 | 206.1096 | 2607 | 0.617 | 206.11 |
| 1,7-dimethylphenanthrene | 1,7-DiMePhe | C16H14 | 206.1096 | 2624 | 0.613 | 206.11 |
| 3,6-dimethylcarbazole | 3,6-DiMeCar | C14H13N | 195.1048 | 2628 | 0.8 | 194.1 |
| fluoranthene-d10 | Flt-d10 | C16D10 | 212.141 | 2640 | 0.891 | 212.14 |
| 1,2-dimethyldibenzothiophene | 1,2-DiMeDBT | C14H12S | 212.066 | 2642 | 0.667 | 197.04 |
| fluoranthene | Flt | C16H10 | 202.0783 | 2646 | 1.08 | 202.08 |
| 2-bromodibenzo[b,d]thiophene | 2-Br-DBT | C12H7BrS | 262.9449 | 2650 | 0.853 | 261.94 |
| 1,8-dimethylphenanthrene | 1,8-DiMePhe | C16H14 | 206.1096 | 2659 | 0.711 | 206.11 |
| 2,4,7-trimethyldibenzothiophene | 2,4,5-TriMeDBT | C15H14S | 226.0817 | 2682 | 0.177 | 226.08 |
| 9-bromoacridine | 9-Br-Acr | C13H8BrN | 256.9837 | 2684 | 0.645 | 255.99 |
| 9-bromoanthracene | 9-Br-Ant | C14H9Br | 255.9888 | 2708 | 0.599 | 255.99 |
| pyrene-d10 | Pyr-d10 | C16D10 | 212.141 | 2709 | 0.822 | 212.14 |
| pyrene | Pyr | C16H10 | 202.0783 | 2714 | 0.82 | 202.08 |
| 4,6-diethyldibenzothiophene | 4,6-DiEtDBT | C16H16S | 240.0973 | 2724 | 2.029 | 225.07 |
| benzo[b]naphtho[2,3-d]furan | BbNF(2,3-d) | C16H10O | 218.0732 | 2758 | 0.44 | 218.07 |
| 3-tert-butyl-9H-carbazole | 3-tBut-Car | C16H17N | 223.1361 | 2770 | 0.192 | 208.11 |
| 1,4-dichloroanthracene | 1,4-Cl2-Ant | C14H8Cl2 | 246.0004 | 2791 | 0.246 | 246 |
| 1,2,6-trimethylphenanthrene | 1,2,6-TriMePhe | C17H16 | 220.1252 | 2807 | 2.1 | 220.12 |
| retene | Ret | C18H18 | 234.1409 | 2819 | 1.887 | 219.12 |
| 5,6-dibromo-1,2-dihydroacenaphthylene | 5,6-Br2-1,2-H2-Acy | C12H8Br2 | 311.9052 | 2836 | 0.605 | 152.06 |
| 2,7-dibromofluoranthene | 2,7-Br2-Flu | C16H8Br2 | 359.8895 | 2860 | 0.446 | 242.98 |
| 1-methylpyrene | 1-MePyr | C17H12 | 216.0939 | 2878 | 0.6 | 215.08 |
| 9-phenylcarbazole | 9-Ph-Car | C18H13N | 243.1048 | 2900 | 0.171 | 243.1 |
| 1,2,6,9-tetramethylphenanthrene | 1,2,6,9-TetMePhe | C18H18 | 234.1409 | 2926 | 2.829 | 234.14 |
| 3,6-dichloro-9H-carbazole | 3,6-Cl2-Car | C12H7Cl2N | 235.9797 | 2945 | 0.4 | 234.99 |
| 4,5-dimethylpyrene | 4,5-DiMePyr | C18H14 | 230.1096 | 2947 | 0.186 | 215.09 |
| 7,8,9,10-tetrahydrobenzo[b]naphtho[2,3-d]thiophene | 7,8,9,10-H4-BbNT(2,3-d) | C16H14S | 238.0817 | 2974 | 0.137 | 210.05 |
| 3-bromofluoranthene | 1-BrFlt | C16H9Br | 280.9888 | 2986 | 0.454 | 279.99 |
| benz[c]acridine | BcAcr | C17H11N | 229.0891 | 2988 | 0.406 | 229.09 |
| benzo[b]naphtho[1,2-d]thiophene | BbNT(1,2-d) | C16H10S | 234.0504 | 2992 | 0.409 | 234.05 |
| benzo[b]naphtho[2,3-d]thiophene | BbNT(2,3-d) | C16H10S | 234.0504 | 3010 | 0.389 | 234.05 |
| 1-propylpyrene | 1-PropPyr | C19H16 | 244.1252 | 3019 | 0.197 | 215.08 |
| 9,10-dibromoanthracene | 9,10-Br2-Ant | C14H8Br2 | 333.8895 | 3022 | 0.378 | 240.17 |

| | | | | | | |
|---|---|---|---|---|---|---|
| chrysene | Chr | C18H12 | 228.0939 | 3027 | 0.356 | 228.09 |
| chrysene-d12 | Chr-d12 | C18D12 | 240.1884 | 3030 | 0.418 | 240.17 |
| 1-bromopyrene | 1-Br-Pyr | C16H9Br | 280.9888 | 3031 | 0.6 | 201.07 |
| 11H-benzo[a]carbazole | 11H-BaCar | C16H11N | 217.0891 | 3034 | 0.541 | 217.09 |
| benzo[a]anthracene | BaAnt | C18H12 | 228.0939 | 3033 | 0.406 | 228.09 |
| 7H-benzo[c]carbazole | 7H-BcCar | C16H11N | 217.0891 | 3085 | 0.801 | 218.98 |
| 6-methylbenzo[b]naphtho[2,1-d]thiophene | 6-MeBbNT(2,1-d) | C17H12S | 248.066 | 3091 | 0.348 | 248.06 |
| 1-butylpyrene | 1-ButPyr | C20H18 | 258.1409 | 3096 | 1.533 | 215.08 |
| 2-phenyldibenzothiophene | 2-Ph-DBT | C18H12S | 260.066 | 3124 | 0.32 | 260.06 |
| 3,6-dibromocarbazole | 3,6-Br2-Car | C12H7Br2N | 323.8787 | 3126 | 0.77 | 324.89 |
| 9-bromo-1,5-chloroanthracene | 9-Br-1,5-Cl-Ant | C14H8BrCl | 290.9497 | 3130 | 0.5 | 242.11 |
| 5-ethylchrysene | 5-EtChr | C20H16 | 256.1252 | 3174 | 0.603 | 241.1 |
| 6-propylchrysene | 6-PropChr | C21H18 | 270.1409 | 3225 | 0.837 | 241.1 |
| benzo[b]fluoranthene-d12 | BbF-d12 | C20D12 | 264.1884 | 3244 | 1.509 | 264.17 |
| benzo[k]fluoranthene-d13 | BkF-d12 | C20D12 | 264.2685 | 3250 | 1.586 | 264.17 |
| benzo[b]fluoranthene | BbF | C20D12 | 264.2685 | 3250 | 1.599 | 252.09 |
| benzo[k]fluoranthene | BkF | C20H12 | 252.0939 | 3255 | 1.63 | 252.09 |
| dinaphtho[1,2;1,2-d]furan | DNF[1,2;1,2-d] | C20H12O | 268.0888 | 3288 | 1.812 | 268.09 |
| 4-bromobenzo[a]anthracene | 4-Br-BaA | C18H11Br | 306 | 3289 | 1.928 | 226.08 |
| 1,6-dibromopyrene | 1,6-Br2-Pyr | C16H8Br2 | 360.0419 | 3287 | 2.064 | 200.06 |
| 6-butylchrysene | 6-ButChr | C22H18 | 282.1409 | 3298 | 1.282 | 241.1 |
| perylene-d12 | Per-d12 | C20D12 | 264.2685 | 3316 | 2.562 | 264.17 |
| perylene | Per | C20H12 | 252.0939 | 3322 | 2.643 | 252.09 |
| benzo[a]pyrene-d12 | BaP-d12 | C20D12 | 264.2685 | 3337 | 2.868 | 264.17 |
| 6-methylbenzo(a)pyrene | 6-MeBaP | C21H14 | 266.1096 | 3517 | 1.757 | 266.11 |
| 7,10-dimethylbenzo(a)pyrene | 7,8-DiMeBaP | C22H16 | 280.1252 | 3610 | 2.46 | 280.12 |
| Indeno(1,2,3-cd)pyrene d12 | IcdP-d12 | C22D12 | 288.2685 | 3637 | 0.24 | 288.17 |
| Indeno(1,2,3-cd)pyrene | IcdP | C22H12 | 276.0939 | 3646 | 0.386 | 276.09 |
| Dibenz(a,h)anthracene d14 | DBahA-d14 | C22D14 | 292.2841 | 3640 | 0.146 | 292.2 |
| Dibenz(a,h)anthracene | DBahA | C22H14 | 278.1096 | 3649 | 0.317 | 278.11 |
| 7H-Dibenzo[c,g]carbazole | 7H-DiBenzCar | C20H13N | 267.1048 | 3709 | 0.197 | 267.1 |
| Benzo(g,h,i)perylene d12 | BghiP-d12 | C22D12 | 288.2685 | 3724 | 1.877 | 288.17 |
| Benzo(g,h,i)perylene | BghiP | C22H12 | 276.0939 | 3733 | 2.049 | 276.09 |

**Table V 4:** The results of a standard processed through the workflow 1

| Peak assignment from ChromaTOF | Peak assignment from WF1 | Confidence | Matched ion counts | RT1 | RT2 | ΔRT1 with library data | ΔRT2 with library data |
|---|---|---|---|---|---|---|---|
| Peak 5 | Nap-d8 | Level 1: Confirmed | 4 | 880 | 2.983 | 0.02 | 0.05 |
| Acetic acid, 2-(2-hydroxy- 1-naphthylmethylamino)- | Nap | Level 1: Confirmed | 4 | 889 | 0.000 | 0.08 | 0.10 |
| Peak 8 | 2,3-Br2-BT | Level 1: Confirmed | 4 | 889 | 0.163 | 0.03 | 0.06 |
| Benzo[b]thiophene | BbT | Level 1: Confirmed | 4 | 904 | 0.197 | 0.05 | 0.06 |
| 05-Phenyl-1,2,3-thiadiazole | BbT | Level 3: Isomer | 3 | 907 | 0.189 | 0.10 | 0.05 |
| Quinoline | Qui | Level 1: Confirmed | 4 | 985 | 0.609 | 0.07 | 0.07 |
| Benzeneacetonitrile, a-acetyl- | Ind | Level 1: Confirmed | 4 | 1087 | 1.089 | 0.05 | 0.10 |
| 2-Methylnaphthalene | 2-MeNap | Level 1: Confirmed | 4 | 1099 | 0.283 | 0.05 | 0.06 |
| Indole | 2-MeQui | Level 1: Confirmed | 4 | 1120 | 0.669 | 0.03 | 0.06 |
| Peak 27 | 4-MeBT | Level 1: Confirmed | 3 | 1132 | 0.583 | 0.05 | 0.07 |
| 1-methylnaphthalene | 1-MeNap | Level 1: Confirmed | 4 | 1135 | 0.429 | 0.05 | 0.04 |
| 3-(4-Methyl-1-quinolinio)-1-propanesulfonate | 3-MeIQui | Level 1: Confirmed | 4 | 1162 | 0.677 | 0.05 | 0.07 |
| Peak 41 | 5-Cl-BT | Level 1: Confirmed | 4 | 1300 | 0.977 | 0.03 | 0.09 |
| Peak 47 | 2,3-DiMeBT | Level 1: Confirmed | 4 | 1378 | 0.737 | 0.03 | 0.08 |
| 6-Chloro-8-quinolinecarboxylic acid | 2-Cl-Qui | Level 1: Confirmed | 4 | 1381 | 1.346 | 0.03 | 0.11 |
| Peak 52 | 1,4-DiMeNap | Level 1: Confirmed | 4 | 1405 | 0.711 | 0.03 | 0.08 |
| Thiazolidin-4-one, 5-(4-bromobenzylideno)-2-thioxo- | 2-Br-BT | Level 1: Confirmed | 4 | 1408 | 1.029 | 0.05 | 0.11 |
| Acenphthalene d8 | Acy-d8 | Level 1: Confirmed | 4 | 1426 | 1.166 | 0.07 | 0.08 |
| 1-acenaphthenol, trifluoroacetate ester | Acy | Level 1: Confirmed | 4 | 1429 | 1.2 | 0.05 | 0.07 |
| Peak 57 | 1,4-DiMeNap | Level 3: Isomer | 3 | 1438 | 0.751 | 0.58 | 0.12 |
| Acenaphthylene d8 | Ace-d10 | Level 1: Confirmed | 4 | 1492 | 1.097 | 0.05 | 0.09 |
| (6Balpha,7beta,11alpha,11aalpha)-11-phenyl-6b,7,11,11a-tetrahydro-7,11-ethenocyclohept(a)acenaphthylen-8-one | Ace | Level 1: Confirmed | 4 | 1504 | 1.131 | 0.05 | 0.09 |
| 2,3-dimethylindole | 2,3-DiMeInd | Level 1: Confirmed | 4 | 1507 | 1.414 | 0 | 0.13 |
| 2,3-Dichlorobenzo[b]thiophene | 2,3-Cl2-BT | Level 1: Confirmed | 4 | 1540 | 0.857 | 0.07 | 0.09 |
| Dibenzofuran | DBF | Level 1: Confirmed | 3 | 1573 | 1.071 | 0.07 | 0.08 |
| 2-bromoquinoline | 2-Br-Qui | Level 1: Confirmed | 4 | 1606 | 1.826 | 0.03 | 0.13 |
| 2-(p-Chlorophenyl)-3-methylbutyronitrile | 5-Cl-Ind | Level 1: Confirmed | 4 | 1624 | 1.911 | 0 | 0.14 |
| 2,3,5-Trimethylnaphthalene | 2,3,5-TriMeNap | Level 1: Confirmed | 4 | 1675 | 0.72 | 0.05 | 0.08 |
| Peak 92 | Flu-d10 | Level 1: Confirmed | 4 | 1708 | 1.234 | 0.07 | 0.09 |
| Peak 93 | Flu-d10 | Candidate | 2 | 1711 | 1.206 | 0.12 | 0.06 |
| 3-bromoisoquinoline | 3-Br-IQui | Level 1: Confirmed | 4 | 1714 | 2.006 | 0.05 | 0.11 |
| 9H-Fluorene, 9-bromo- | Flu | Level 1: Confirmed | 4 | 1720 | 1.26 | 0.07 | 0.09 |
| 2,3,4-Trimethyl-Benzothiophene | 2,3,4-TriMeBT | Level 1: Confirmed | 4 | 1726 | 1.046 | 0.07 | 0.08 |
| Xanthene-9-carboxylic acid | 2-MeDBF | Level 1: Confirmed | 4 | 1828 | 1.054 | 0.05 | 0.09 |
| 5-Bromoindole | 5-Br-Ind | Level 1: Confirmed | 4 | 1843 | 2.349 | 0.02 | 0.11 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 9-n-propylfluorene | 9-EtFlu | Level 1: Confirmed | 4 | 1921 | 1.054 | 0.07 | 0.09 |
| Peak 118 | 1,4,6,7-TetraMeNap | Level 1: Confirmed | 4 | 1936 | 0.746 | 0.05 | 0.09 |
| (3H)Benzo[c]pyrrole, 3-methyl-3-phenyl- | 1-MeFlu | Level 1: Confirmed | 4 | 1981 | 1.311 | 0.07 | 0.11 |
| Peak 123 | 1-MeFlu | Candidate | 2 | 1987 | 1.258 | 0.17 | 0.05 |
| 4-(Naphthalen-1-yl)-1,2,3-thiadiazole | DBT | Level 1: Confirmed | 4 | 2080 | 1.937 | 0.03 | 0.13 |
| 9-Ethylfluorene | 9-PropFlu | Level 1: Confirmed | 4 | 2113 | 0.969 | 0.07 | 0.09 |
| Peak 133 | 5-Br-Ace | Level 3: Isomer | 3 | 2116 | 1.71 | 1.42 | 0.04 |
| Peak 134 | Phe-d10 | Level 1: Confirmed | 4 | 2128 | 1.937 | 0.02 | 0.1 |
| Phenanthrene | Phe | Level 1: Confirmed | 4 | 2137 | 1.971 | 0.03 | 0.14 |
| Anthracene d10 | Ant-d10 | Level 1: Confirmed | 4 | 2149 | 1.911 | 0.03 | 0.13 |
| 9,10-Dihydro-9-(2-oxocyclopentyl)acridine | BhQui | Level 1: Confirmed | 4 | 2155 | 2.289 | 0.03 | 0.15 |
| Dibenzo[b,e]7,8-diazabicyclo[2.2.2]octa-2,5-diene | Ant | Level 1: Confirmed | 4 | 2158 | 1.937 | 0.07 | 0.13 |
| Peak 141 | 1,2,3,4-TetHCar | Level 1: Confirmed | 4 | 2179 | 2.229 | 0.05 | 0.15 |
| Peak 145 | 5-Br-Ace | Level 1: Confirmed | 4 | 2203 | 1.877 | 0.03 | 0.13 |
| 2-(1H-Benzo[g]indol-3-yl)-1-methylethylamine | 9-MeCar | Level 1: Confirmed | 4 | 2221 | 2.134 | 0 | 0 |
| Carbazole | Car | Level 1: Confirmed | 4 | 2245 | 2.7 | 0 | 0.2 |
| Peak 150 | 2-MeDBT | Level 1: Confirmed | 4 | 2284 | 1.783 | 0.03 | 0.12 |
| 9-butylfluorene | 9-ButFlu | Level 1: Confirmed | 4 | 2296 | 0.909 | 0.07 | 0.06 |
| Peak 154 | 3-MeDBT | Level 1: Confirmed | 4 | 2323 | 1.749 | 0.03 | 0.12 |
| 3-Methylphenanthrene | 3-MePhe | Level 1: Confirmed | 4 | 2365 | 1.774 | 0.03 | 0.12 |
| 9-Phenanthrenol, 9,10-dihydro-10-[(triphenylphosphoranylidene)amino]-, trans-(.+-.)- | 3-MeBfQui | Level 1: Confirmed | 4 | 2365 | 2.203 | 0.02 | 0.17 |
| 9-methylphenanthrene | 2-MePhe | Level 1: Confirmed | 4 | 2377 | 1.843 | 0.03 | 0.09 |
| Peak 161 | 9-Br-Flu | Level 1: Confirmed | 4 | 2389 | 1.929 | 0.03 | 0.12 |
| 5-Acetyl-5H-dibenz[b,f]azepine | 2-MeAcr | Level 1: Confirmed | 4 | 2419 | 2.134 | 0.02 | 0.15 |
| 9-Hydroxymethyl-9,10-dihydro-2-chloro acridine | 9-Cl-Acr | Level 1: Confirmed | 4 | 2503 | 1.774 | 0.02 | 0.17 |
| Peak 173 | 2,4-DiMeDBT | Level 1: Confirmed | 4 | 2545 | 0.977 | 0 | 0 |
| Benzo[h]quinoline, 2,4-dimethyl- | 2,4-DiMeBhQ | Level 1: Confirmed | 4 | 2557 | 1.209 | 0.02 | 0.21 |
| Peak 175 | 1-Cl-Ant | Level 1: Confirmed | 4 | 2557 | 1.311 | 0.02 | 0.22 |
| Anthracene, 9-ethyl-9,10-dihydro-10-hydroxy- | 3,6-DiMePhe | Level 1: Confirmed | 4 | 2566 | 0.874 | 0.02 | 0.18 |
| Peak 180 | 2,6-DiMePhe | Level 1: Confirmed | 4 | 2581 | 0.831 | 0.03 | 0.17 |
| Peak 181 | 2,7-Cl2-Flu | Level 1: Confirmed | 4 | 2593 | 1.011 | 0.02 | 0.2 |
| Peak 182 | 1,2-Br2-Acy | Level 1: Confirmed | 4 | 2596 | 1.209 | 0 | 0.21 |
| Peak 184 | 4-PropDBT | Level 1: Confirmed | 4 | 2608 | 0.634 | 0.03 | 0.15 |
| Peak 186 | 4-PropDBT | Level 3: Isomer | 3 | 2623 | 0.625 | 0.28 | 0.15 |
| 5H-Dibenz[b,f]azepine-5-carboxamide, 10,11-dihydro- | 3,6-DiMeCar | Level 1: Confirmed | 4 | 2629 | 1.003 | 0.02 | 0.2 |
| Peak 190 | 1,7-DiMePhe | Level 1: Confirmed | 3 | 2629 | 0.803 | 0.08 | 0.19 |
| Peak 193 | Flt-d10 | Level 1: Confirmed | 4 | 2641 | 1.089 | 0.02 | 0.2 |
| Peak 194 | 1,2-DiMeDBT | Level 1: Confirmed | 4 | 2644 | 0.84 | 0.03 | 0.17 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1,2-dimethyldibenzothiophene | 1,2-DiMeDBT | Level 2: Probable | 2 | 2647 | 0.789 | 0.08 | 0.12 |
| Fluoranthene | 2-Br-DBT | Level 2: Probable | 2 | 2650 | 1.011 | 0 | 0.16 |
| Peak 205 | 2,4,5-TriMeDBT | Level 1: Confirmed | 4 | 2683 | 0.334 | 0.02 | 0.16 |
| Peak 207 | 9-Br-Acr | Level 1: Confirmed | 4 | 2686 | 0.814 | 0.03 | 0.17 |
| 9-Bromoanthracene | 9-Br-Ant | Level 1: Confirmed | 4 | 2710 | 0.763 | 0.03 | 0.16 |
| Pyrene d10 | Pyr-d10 | Level 1: Confirmed | 4 | 2710 | 1.011 | 0.02 | 0.19 |
| Peak 213 | Pyr | Level 1: Confirmed | 4 | 2716 | 0.994 | 0.03 | 0.17 |
| Peak 214 | Pyr | Candidate | 2 | 2722 | 0.765 | 0.13 | 0.05 |
| Peak 215 | 4,6-DiEtDBT | Level 3: Isomer | 4 | 2728 | 0.154 | 0.07 | 1.88 |
| Peak 216 | BbNF(2,3-d) | Level 1: Confirmed | 4 | 2758 | 0.66 | 0 | 0.22 |
| Peak 219 | 3-tBut-Car | Level 1: Confirmed | 4 | 2770 | 0.437 | 0 | 0.25 |
| Ethene, 1-chloro-1,2-bis(4-chlorophenyl)- | 1,4-Cl2-Ant | Level 1: Confirmed | 4 | 2791 | 0.506 | 0 | 0.26 |
| Peak 223 | 1,4-Cl2-Ant | Level 3: Isomer | 3 | 2806 | 0.531 | 0.25 | 0.29 |
| Phenanthrene, 1,2,6-trimethyl- | 1,2,6-TriMePhe | Level 3: Isomer | 4 | 2809 | 0.343 | 0.03 | 1.76 |
| Peak 225 | Ret | Level 3: Isomer | 4 | 2821 | 0.103 | 0.03 | 1.78 |
| Peak 228 | 5,6-Br2-1,2-H2-Acy | Level 1: Confirmed | 4 | 2836 | 0.889 | 0 | 0.28 |
| 2,7-dibromofluoranthene | 2,7-Br2-Flu | Level 1: Confirmed | 4 | 2860 | 0.566 | 0 | 0.12 |
| Pyrene, 1-(chloromethyl)- | 1-MePyr | Level 1: Confirmed | 4 | 2878 | 0.6 | 0 | 0 |
| Peak 242 | 9-Ph-Car | Level 1: Confirmed | 4 | 2902 | 0.3 | 0.03 | 0.13 |
| Peak 247 | 1,2,6,9-TetMePhe | Level 1: Confirmed | 4 | 2926 | 2.94 | 0 | 0.11 |
| Peak 250 | 3,6-Cl2-Car | Level 1: Confirmed | 4 | 2944 | 0.557 | 0.02 | 0.16 |
| Pyrene, 4,5-dimethyl- | 4,5-DiMePyr | Level 1: Confirmed | 4 | 2947 | 0.317 | 0 | 0.13 |
| Peak 258 | 7,8,9,10-H4-BbNT(2,3-d) | Level 1: Confirmed | 4 | 2974 | 0.257 | 0 | 0.12 |
| Peak 265 | 1-BrFlt | Level 1: Confirmed | 4 | 2986 | 0.454 | 0 | 0 |
| Peak 266 | BcAcr | Level 1: Confirmed | 4 | 2989 | 0.54 | 0.02 | 0.13 |
| Peak 268 | BbNT(1,2-d) | Level 1: Confirmed | 4 | 2992 | 0.549 | 0 | 0.14 |
| Benzo[b]naphtho[2,3-d]thiophene | BbNT(2,3-d) | Level 1: Confirmed | 4 | 3010 | 0.531 | 0 | 0.14 |
| 1-n-propylpyrene | 1-PropPyr | Level 1: Confirmed | 3 | 3019 | 0.197 | 0 | 0 |
| Peak 278 | 9,10-Br2-Ant | Level 1: Confirmed | 4 | 3022 | 0.489 | 0 | 0.11 |
| Chrysene | Chr | Level 1: Confirmed | 4 | 3028 | 0.48 | 0.02 | 0.12 |
| 1-Bromopyrene | 1-Br-Pyr | Level 1: Confirmed | 4 | 3031 | 0.6 | 0 | 0 |
| Chrysene-D12 | Chr-d12 | Level 1: Confirmed | 4 | 3031 | 0.549 | 0.02 | 0.13 |
| 11H-Benzo[a]carbazole | 11H-BaCar | Level 1: Confirmed | 4 | 3034 | 0.688 | 0 | 0.15 |
| 5,12-ethanonaphthacene, 5,12-dihydro- | BaAnt | Level 1: Confirmed | 4 | 3037 | 0.557 | 0.07 | 0.15 |
| Isophthalic acid, 2-bromo-4-fluorophenyl pentyl ester | 7H-BcCar | Level 3: Isomer | 4 | 3043 | 1.183 | 0.7 | 0.38 |
| Terephthalic acid, pentyl tridec-2-yn-1-yl ester | 7H-BcCar | Level 3: Isomer | 4 | 3064 | 1.131 | 0.35 | 0.33 |
| Peak 308 | 7H-BcCar | Level 1: Confirmed | 4 | 3085 | 0.891 | 0 | 0.09 |
| Peak 312 | 6-MeBbNT(2,1-d) | Level 1: Confirmed | 4 | 3091 | 0.471 | 0 | 0.12 |
| Peak 327 | 2-Ph-DBT | Level 1: Confirmed | 4 | 3124 | 0.429 | 0 | 0.11 |
| Peak 328 | 3,6-Br2-Car | Level 1: Confirmed | 4 | 3124 | 0.871 | 0.03 | 0.1 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 9-Bromo-1,5-chloroanthracene | 9-Br-1,5-Cl-Ant | Level 1: Confirmed | 4 | 3130 | 0.609 | 0 | 0.11 |
| Chrysene, 5-ethyl- | 5-EtChr | Level 1: Confirmed | 4 | 3175 | 0.677 | 0.02 | 0.07 |
| Chrysene, 6-propyl- | 6-PropChr | Level 1: Confirmed | 4 | 3226 | 0.891 | 0.02 | 0.05 |
| Peak 430 | BbF-d12 | Level 1: Confirmed | 4 | 3244 | 1.534 | 0 | 0.03 |
| Benzo[k]fluoranthene-d12 | BkF-d12 | Level 1: Confirmed | 4 | 3247 | 1.611 | 0.05 | 0.02 |
| Peak 436 | BbF | Level 1: Confirmed | 4 | 3250 | 1.646 | 0 | 0.05 |
| Benzo[k]fluoranthene | BkF | Level 1: Confirmed | 4 | 3256 | 1.68 | 0.02 | 0.05 |
| Benzo[b]fluoranthene | BkF | Level 2: Probable | 2 | 3259 | 1.674 | 0.07 | 0.04 |
| Acetic acid, perylen-3-yl ester | DNF[1,2;1,2-d] | Level 1: Confirmed | 4 | 3289 | 1.869 | 0.02 | 0.06 |
| Peak 480 | 4-Br-BaA | Level 1: Confirmed | 4 | 3289 | 1.99 | 0 | 0.06 |
| Peak 481 | 1,6-Br2-Pyr | Level 1: Confirmed | 4 | 3289 | 2.097 | 0.03 | 0.03 |
| 6-Butylchrysene | 6-ButChr | Level 1: Confirmed | 4 | 3298 | 1.329 | 0 | 0.05 |
| Perylene-D12 | Per-d12 | Level 1: Confirmed | 4 | 3316 | 2.595 | 0 | 0.03 |
| Perylene | Per | Level 1: Confirmed | 4 | 3322 | 2.674 | 0 | 0.03 |
| Peak 520 | BaP-d12 | Level 2: Probable | 2 | 3334 | 2.955 | 0.05 | 0.09 |
| Peak 524 | BaP-d12 | Level 1: Confirmed | 4 | 3337 | 2.889 | 0 | 0.02 |
| Peak 721 | 7,8-DiMeBaP | Level 1: Confirmed | 3 | 3610 | 2.642 | 0 | 0.18 |
| Peak 724 | 7,8-DiMeBaP | Level 1: Confirmed | 3 | 3613 | 2.65 | 0.05 | 0.19 |
| Peak 750 | IcdP | Level 2: Probable | 2 | 3643 | 0.577 | 0.05 | 0.19 |
| Dibenz[a,j]anthracene | DBahA | Level 1: Confirmed | 3 | 3649 | 0.534 | 0 | 0.22 |
| Peak 826 | BghiP | Level 1: Confirmed | 3 | 3733 | 2.306 | 0 | 0.26 |

**Table V 5:** Aromaticity indices for the targeted analytes

| Acronym | Formula | Exact Mass | Estimated number of carbons | Mass defect | Aromaticity Index | Class |
|---|---|---|---|---|---|---|
| Ace-d10 | C12D10 | 164.141 | 12.6262 | 0.141 | 0.0112 | Parent/Alkyl-PAH |
| Nap-d8 | C10D8 | 136.1128 | 10.4702 | 0.1128 | 0.0108 | Parent/Alkyl-PAH |
| Flu-d10 | C13D10 | 176.141 | 13.5493 | 0.141 | 0.0104 | Parent/Alkyl-PAH |
| Phe-d10 | C14D10 | 188.141 | 14.4724 | 0.141 | 0.0097 | Parent/Alkyl-PAH |
| Ant-d10 | C14D10 | 188.141 | 14.4724 | 0.141 | 0.0097 | Parent/Alkyl-PAH |
| Acy-d8 | C12D8 | 160.1128 | 12.3164 | 0.1128 | 0.0092 | Parent/Alkyl-PAH |
| Chr-d12 | C18D12 | 240.1692 | 18.4746 | 0.1692 | 0.0092 | Parent/Alkyl-PAH |
| 1,4,6,7-TetraMeNap | C14H16 | 184.1252 | 14.1635 | 0.1252 | 0.0088 | Parent/Alkyl-PAH |
| DBahA-d14 | C22D14 | 292.1974 | 22.4767 | 0.1974 | 0.0088 | Parent/Alkyl-PAH |
| Flt-d10 | C16D10 | 212.141 | 16.3185 | 0.141 | 0.0086 | Parent/Alkyl-PAH |
| Pyr-d10 | C16D10 | 212.141 | 16.3185 | 0.141 | 0.0086 | Parent/Alkyl-PAH |
| 2,3,5-TriMeNap | C13H14 | 170.1096 | 13.0854 | 0.1096 | 0.0084 | Parent/Alkyl-PAH |
| BbF-d12 | C20D12 | 264.1692 | 20.3207 | 0.1692 | 0.0083 | Parent/Alkyl-PAH |
| BkF-d12 | C20D12 | 264.1692 | 20.3207 | 0.1692 | 0.0083 | Parent/Alkyl-PAH |
| BbF | C20D12 | 264.1692 | 20.3207 | 0.1692 | 0.0083 | Parent/Alkyl-PAH |
| Per-d12 | C20D12 | 264.1692 | 20.3207 | 0.1692 | 0.0083 | Parent/Alkyl-PAH |
| BaP-d12 | C20D12 | 264.1692 | 20.3207 | 0.1692 | 0.0083 | Parent/Alkyl-PAH |
| 9-ButFlu | C17H18 | 222.1409 | 17.0878 | 0.1409 | 0.0082 | Parent/Alkyl-PAH |
| 7H-BcCar-d12 | C20D12N | 278.1723 | 21.3979 | 0.1723 | 0.0081 | NPAC (Nitrogen) |
| 2,3-DiMeInd | C10H11N | 145.0891 | 11.1607 | 0.0891 | 0.008 | NPAC (Nitrogen) |
| 1,2,3,4-TetHCar | C12H13N | 171.1048 | 13.1619 | 0.1048 | 0.008 | NPAC (Nitrogen) |
| 3-tBut-Car | C16H17N | 223.1361 | 17.1643 | 0.1361 | 0.0079 | NPAC (Nitrogen) |
| 9-PropFlu | C16H16 | 208.1252 | 16.0096 | 0.1252 | 0.0078 | Parent/Alkyl-PAH |
| Ret | C18H18 | 234.1409 | 18.0108 | 0.1409 | 0.0078 | Parent/Alkyl-PAH |
| 1,2,6,9-TetMePhe | C18H18 | 234.1409 | 18.0108 | 0.1409 | 0.0078 | Parent/Alkyl-PAH |
| 1,4-DiMeNap | C12H12 | 156.0939 | 12.0072 | 0.0939 | 0.0078 | Parent/Alkyl-PAH |
| IcdP-d12 | C22D12 | 288.1692 | 22.1669 | 0.1692 | 0.0076 | Parent/Alkyl-PAH |
| BghiP-d12 | C22D12 | 288.1692 | 22.1669 | 0.1692 | 0.0076 | Parent/Alkyl-PAH |
| 1,2,6-TriMePhe | C17H16 | 220.1252 | 16.9327 | 0.1252 | 0.0074 | Parent/Alkyl-PAH |
| 9-EtFlu | C15H14 | 194.1096 | 14.9315 | 0.1096 | 0.0073 | Parent/Alkyl-PAH |
| 2-MeNap | C11H10 | 142.0783 | 10.9291 | 0.0782 | 0.0072 | Parent/Alkyl-PAH |
| 1-MeNap | C11H10 | 142.0783 | 10.9291 | 0.0782 | 0.0072 | Parent/Alkyl-PAH |
| 1-ButPyr | C20H18 | 258.1409 | 19.857 | 0.1409 | 0.0071 | Parent/Alkyl-PAH |
| 3,6-DiMeCar | C14H13N | 195.1048 | 15.0081 | 0.1048 | 0.007 | NPAC (Nitrogen) |
| 3,6-DiMePhe | C16H14 | 206.1096 | 15.8546 | 0.1096 | 0.0069 | Parent/Alkyl-PAH |
| 2,6-DiMePhe | C16H14 | 206.1096 | 15.8546 | 0.1096 | 0.0069 | Parent/Alkyl-PAH |
| 1,3-DiMePhe | C16H14 | 206.1096 | 15.8546 | 0.1096 | 0.0069 | Parent/Alkyl-PAH |
| 1,7-DiMePhe | C16H14 | 206.1096 | 15.8546 | 0.1096 | 0.0069 | Parent/Alkyl-PAH |
| 1,8-DiMePhe | C16H14 | 206.1096 | 15.8546 | 0.1096 | 0.0069 | Parent/Alkyl-PAH |
| 6-PropChr | C21H18 | 270.1409 | 20.7801 | 0.1409 | 0.0068 | Parent/Alkyl-PAH |
| 1-MeFlu | C14H12 | 180.0939 | 13.8534 | 0.0939 | 0.0068 | Parent/Alkyl-PAH |
| 1-MeFlu | C14H12 | 180.0939 | 13.8534 | 0.0939 | 0.0068 | Parent/Alkyl-PAH |
| 2-MeQui | C10H9N | 143.0735 | 11.0057 | 0.0735 | 0.0067 | NPAC (Nitrogen) |
| 3-MeIQui | C10H9N | 143.0735 | 11.0057 | 0.0735 | 0.0067 | NPAC (Nitrogen) |
| 1-PropPyr | C19H16 | 244.1252 | 18.7789 | 0.1252 | 0.0067 | Parent/Alkyl-PAH |
| Ace | C12H10 | 154.0783 | 11.8522 | 0.0782 | 0.0066 | Parent/Alkyl-PAH |
| 2,4-DiMeBhQ | C15H13N | 207.1048 | 15.9311 | 0.1048 | 0.0066 | NPAC (Nitrogen) |
| 6-ButChr | C22H18 | 282.1409 | 21.7031 | 0.1409 | 0.0065 | Parent/Alkyl-PAH |
| Ind | C8H7N | 117.0578 | 9.0044 | 0.0578 | 0.0064 | NPAC (Nitrogen) |
| 9-MeCar | C13H11N | 181.0891 | 13.9299 | 0.0891 | 0.0064 | NPAC (Nitrogen) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Nap | C10H8 | 128.0626 | 9.851 | 0.0626 | 0.0064 | Parent/Alkyl-PAH |
| 5-EtChr | C20H16 | 256.1252 | 19.7019 | 0.1252 | 0.0064 | Parent/Alkyl-PAH |
| 3-MePhe | C15H12 | 192.0939 | 14.7765 | 0.0939 | 0.0064 | Parent/Alkyl-PAH |
| 2-MePhe | C15H12 | 192.0939 | 14.7765 | 0.0939 | 0.0064 | Parent/Alkyl-PAH |
| 9-MePhe | C15H12 | 192.0939 | 14.7765 | 0.0939 | 0.0064 | Parent/Alkyl-PAH |
| 1-MePhe | C15H12 | 192.0939 | 14.7765 | 0.0939 | 0.0064 | Parent/Alkyl-PAH |
| 4,5-DiMePyr | C18H14 | 230.1096 | 17.7007 | 0.1096 | 0.0062 | Parent/Alkyl-PAH |
| Flu | C13H10 | 166.0783 | 12.7753 | 0.0782 | 0.0061 | Parent/Alkyl-PAH |
| 3-MeBfQui | C14H11N | 193.0891 | 14.853 | 0.0891 | 0.006 | NPAC (Nitrogen) |
| 2-MeAcr | C14H11N | 193.0891 | 14.853 | 0.0891 | 0.006 | NPAC (Nitrogen) |
| Qui | C9H7N | 129.0578 | 9.9275 | 0.0578 | 0.0058 | NPAC (Nitrogen) |
| IQui | C9H7N | 129.0578 | 9.9275 | 0.0578 | 0.0058 | NPAC (Nitrogen) |
| 7,10-DiMeBaP | C22H16 | 280.1252 | 21.5481 | 0.1252 | 0.0058 | Parent/Alkyl-PAH |
| Car | C12H9N | 167.0735 | 12.8518 | 0.0735 | 0.0057 | NPAC (Nitrogen) |
| Phe | C14H10 | 178.0783 | 13.6983 | 0.0782 | 0.0057 | Parent/Alkyl-PAH |
| Ant | C14H10 | 178.0783 | 13.6983 | 0.0782 | 0.0057 | Parent/Alkyl-PAH |
| 1-MePyr | C17H12 | 216.0939 | 16.6226 | 0.0939 | 0.0056 | Parent/Alkyl-PAH |
| 9-Ph-Car | C18H13N | 243.1048 | 18.7004 | 0.1048 | 0.0056 | NPAC (Nitrogen) |
| 5-MeBaP | C21H14 | 266.1096 | 20.47 | 0.1096 | 0.0054 | Parent/Alkyl-PAH |
| Acy | C12H8 | 152.0626 | 11.6971 | 0.0626 | 0.0054 | Parent/Alkyl-PAH |
| Chr | C18H12 | 228.0939 | 17.5457 | 0.0939 | 0.0054 | Parent/Alkyl-PAH |
| BaAnt | C18H12 | 228.0939 | 17.5457 | 0.0939 | 0.0054 | Parent/Alkyl-PAH |
| 11H-BaCar | C16H11N | 217.0891 | 16.6992 | 0.0891 | 0.0053 | NPAC (Nitrogen) |
| 7H-BcCar | C16H11N | 217.0891 | 16.6992 | 0.0891 | 0.0053 | NPAC (Nitrogen) |
| BhQui | C13H9N | 179.0735 | 13.7749 | 0.0735 | 0.0053 | NPAC (Nitrogen) |
| Acr | C13H9N | 179.0735 | 13.7749 | 0.0735 | 0.0053 | NPAC (Nitrogen) |
| 4,6-DiEtDBT | C16H16S | 240.0973 | 18.469 | 0.0973 | 0.0053 | PASH (Sulfur) |
| 2-MeDBF | C13H10O | 182.0732 | 14.0056 | 0.0732 | 0.0052 | OPAC (Oxygen) |
| DBahA | C22H14 | 278.1096 | 21.393 | 0.1096 | 0.0051 | Parent/Alkyl-PAH |
| 13(H)-DBaiCar | C20H13N | 267.1048 | 20.5465 | 0.1048 | 0.0051 | NPAC (Nitrogen) |
| BcAcr | C17H11N | 229.0891 | 17.6222 | 0.0891 | 0.0051 | NPAC (Nitrogen) |
| Flt | C16H10 | 202.0783 | 15.5445 | 0.0782 | 0.005 | Parent/Alkyl-PAH |
| Pyr | C16H10 | 202.0783 | 15.5445 | 0.0782 | 0.005 | Parent/Alkyl-PAH |
| DBchAcr | C21H13N | 279.1048 | 21.4696 | 0.1048 | 0.0049 | NPAC (Nitrogen) |
| 2,3,4-TriMeBT | C11H12S | 176.066 | 13.5435 | 0.066 | 0.0049 | PASH (Sulfur) |
| BkF | C20H12 | 252.0939 | 19.3918 | 0.0939 | 0.0048 | Parent/Alkyl-PAH |
| Per | C20H12 | 252.0939 | 19.3918 | 0.0939 | 0.0048 | Parent/Alkyl-PAH |
| 4-PropDBT | C15H14S | 226.0816 | 17.3909 | 0.0816 | 0.0047 | PASH (Sulfur) |
| 2,4,5-TriMeDBT | C15H14S | 226.0816 | 17.3909 | 0.0816 | 0.0047 | PASH (Sulfur) |
| 7,8,9,10-H4-BbNT(2,3-d) | C16H14S | 238.0816 | 18.314 | 0.0816 | 0.0045 | PASH (Sulfur) |
| DBF | C12H8O | 168.0575 | 12.9275 | 0.0575 | 0.0044 | OPAC (Oxygen) |
| IcdP | C22H12 | 276.0939 | 21.238 | 0.0939 | 0.0044 | Parent/Alkyl-PAH |
| BghiP | C22H12 | 276.0939 | 21.238 | 0.0939 | 0.0044 | Parent/Alkyl-PAH |
| BbNF(2,3-d) | C16H10O | 218.0732 | 16.7749 | 0.0732 | 0.0044 | OPAC (Oxygen) |
| DNF[1,2;1,2-d] | C20H12O | 268.0888 | 20.6222 | 0.0888 | 0.0043 | OPAC (Oxygen) |
| 2,4-DiMeDBT | C14H12S | 212.066 | 16.3128 | 0.066 | 0.004 | PASH (Sulfur) |
| 1,2-DiMeDBT | C14H12S | 212.066 | 16.3128 | 0.066 | 0.004 | PASH (Sulfur) |
| 2,3-DiMeBT | C10H10S | 162.0503 | 12.4654 | 0.0503 | 0.004 | PASH (Sulfur) |
| 6-MeBbNT(2,1-d) | C17H12S | 248.066 | 19.082 | 0.066 | 0.0035 | PASH (Sulfur) |
| 2-MeDBT | C13H10S | 198.0503 | 15.2346 | 0.0503 | 0.0033 | PASH (Sulfur) |
| 3-MeDBT | C13H10S | 198.0503 | 15.2346 | 0.0503 | 0.0033 | PASH (Sulfur) |
| 2-Ph-DBT | C18H12S | 260.066 | 20.0051 | 0.066 | 0.0033 | PASH (Sulfur) |
| 4-MeBT | C9H8S | 148.0347 | 11.3873 | 0.0347 | 0.003 | PASH (Sulfur) |

| | | | | | | |
|---|---|---|---|---|---|---|
| BbNT(1,2-d) | C16H10S | 234.0503 | 18.0039 | 0.0503 | 0.0028 | PASH (Sulfur) |
| BbNT(2,3-d) | C16H10S | 234.0503 | 18.0039 | 0.0503 | 0.0028 | PASH (Sulfur) |
| 9-Cl-Flu | C13H9Cl | 200.0393 | 15.3876 | 0.0393 | 0.0026 | Halogenated-PAC |
| DBT | C12H8S | 184.0347 | 14.1565 | 0.0347 | 0.0024 | PASH (Sulfur) |
| 1-Cl-Ant | C14H9Cl | 212.0393 | 16.3107 | 0.0393 | 0.0024 | Halogenated-PAC |
| 9-Cl-Acr | C13H8ClN | 213.0345 | 16.3873 | 0.0345 | 0.0021 | Halogenated-PAC |
| BbT | C8H6S | 134.019 | 10.3092 | 0.019 | 0.0018 | PASH (Sulfur) |
| 5-Cl-Ind | C8H6ClN | 151.0189 | 11.6168 | 0.0189 | 0.0016 | Halogenated-PAC |
| 2-Cl-Qui | C9H6ClN | 163.0189 | 12.5399 | 0.0189 | 0.0015 | Halogenated-PAC |
| 4-Br-BaA | C18H11Br | 306.0044 | 23.5388 | 0.0044 | 0.0002 | Halogenated-PAC |
| 2,7-Cl2-Flu | C13H8Cl2 | 234.0003 | 18 | 0.0003 | 0 | Halogenated-PAC |
| 1,4-Cl2-Ant | C14H8Cl2 | 246.0003 | 18.9231 | 0.0003 | 0 | Halogenated-PAC |
| 3,6-Cl2-Car | C12H7Cl2N | 234.9956 | 18.0766 | -0.0044 | -0.0002 | Halogenated-PAC |
| 1-BrFlt | C16H9Br | 279.9888 | 21.5376 | -0.0112 | -0.0005 | Halogenated-PAC |
| 1-Br-Pyr | C16H9Br | 279.9888 | 21.5376 | -0.0112 | -0.0005 | Halogenated-PAC |
| 9-Br-Ant | C14H9Br | 255.9888 | 19.6914 | -0.0112 | -0.0006 | Halogenated-PAC |
| 9-Br-Flu | C13H9Br | 243.9888 | 18.7684 | -0.0112 | -0.0006 | Halogenated-PAC |
| 5-Br-Ace | C12H9Br | 231.9888 | 17.8453 | -0.0112 | -0.0006 | Halogenated-PAC |
| 9-Br-Acr | C13H8BrN | 256.984 | 19.768 | -0.016 | -0.0008 | Halogenated-PAC |
| 5-Cl-BT | C8H5ClS | 167.98 | 12.9215 | -0.02 | -0.0015 | Halogenated-PAC |
| 2-Br-Qui | C9H6BrN | 206.9684 | 15.9206 | -0.0316 | -0.002 | Halogenated-PAC |
| 3-Br-IQui | C9H6BrN | 206.9684 | 15.9206 | -0.0316 | -0.002 | Halogenated-PAC |
| 5-Br-Ind | C8H6BrN | 194.9684 | 14.9976 | -0.0316 | -0.0021 | Halogenated-PAC |
| 9-Br-1,5-Cl-Ant | C14H8BrCl | 289.9498 | 22.3038 | -0.0502 | -0.0023 | Halogenated-PAC |
| 2-Br-DBT | C12H7BrS | 261.9452 | 20.1496 | -0.0548 | -0.0027 | Halogenated-PAC |
| 1,5,9,10-Cl4-Ant | C14H6Cl4 | 313.9224 | 24.1479 | -0.0776 | -0.0032 | Halogenated-PAC |
| 2,7-Br2-Flu | C16H8Br2 | 357.8993 | 27.5307 | -0.1007 | -0.0037 | Halogenated-PAC |
| 1,6-Br2-Pyr | C16H8Br2 | 357.8993 | 27.5307 | -0.1007 | -0.0037 | Halogenated-PAC |
| 2,3-Cl2-BT | C8H4Cl2S | 201.9411 | 15.5339 | -0.0589 | -0.0038 | Halogenated-PAC |
| 9,10-Br2-Ant | C14H8Br2 | 333.8993 | 25.6846 | -0.1007 | -0.0039 | Halogenated-PAC |
| 5,6-Br2-1,2-H2-Acy | C12H8Br2 | 309.8993 | 23.8384 | -0.1007 | -0.0042 | Halogenated-PAC |
| 3,6-Br2-Car | C12H7Br2N | 322.8945 | 24.838 | -0.1055 | -0.0042 | Halogenated-PAC |
| 2-Br-BT | C8H5BrS | 211.9295 | 16.3023 | -0.0705 | -0.0043 | Halogenated-PAC |
| 1,2-Br2-Acy | C12H6Br2 | 307.8836 | 23.6834 | -0.1164 | -0.0049 | Halogenated-PAC |
| 2,3-Br2-BT | C8H4Br2S | 289.84 | 22.2954 | -0.16 | -0.0072 | Halogenated-PAC |
| 2,3-Br2-BT | C4H2Br2S | 239.8244 | 18.448 | -0.1756 | -0.0095 | Halogenated-PAC |

**Table V 6:** The results obtained for benchmarking step of aromaticity index filtering

| Compound category | n | Mean AI | Range | AI behaviour |
|---|---|---|---|---|
| Homocyclic PACs | 84 | 0.0062 | 0.0036 – 0.0088 | Narrow positive band |
| Heterocyclic PACs (N/S/O) | 95 | 0.0039 | 0.0004 – 0.0064 | Narrow positive band |
| Halogenated aromatics | 70 | +0.0007 | −0.0086 – 0.0325 | 83% retained; PBDEs and polybromo non PACs were excluded |
| Halogenated aliphatics | 26 | −0.0031 | −0.0112 – 0.0397 | Largely retained (AI limitation) |
| Non-PAC background | 244 | 0.0099 | −0.0039 – 0.0157 | Saturated aliphatics above band |

**Table V 7:** Possible halogenated PACs that were incorporated in workflow 2 and their exact masses

| Possible compounds | Formula | Exact Mass |
|---|---|---|
| Naphthalene-Mono-Cl | C10ClH7 | 162.0236 |
| Naphthalene-Di-Cl | C10Cl2H6 | 195.9847 |
| Naphthalene-Tri-Cl | C10Cl3H5 | 229.9457 |
| Naphthalene-Tetra-Cl | C10Cl4H4 | 263.9067 |
| Naphthalene-Mono-Br | C10BrH7 | 205.9731 |
| Naphthalene-Di-Br | C10Br2H6 | 283.8836 |
| Naphthalene-Tri-Br | C10Br3H5 | 361.7941 |
| Naphthalene-Tetra-Br | C10Br4H4 | 439.7047 |
| Naphthalene-Mono-Cl-Mono-Br | C10BrClH6 | 239.9341 |
| Naphthalene-Di-Cl-Mono-Br | C10BrCl2H5 | 273.8952 |
| Naphthalene-Mono-Cl-Di-Br | C10Br2ClH5 | 317.8447 |
| Naphthalene-Di-Cl-Di-Br | C10Br2Cl2H4 | 351.8057 |
| Acenaphthylene-Mono-Cl | C12ClH7 | 186.0236 |
| Acenaphthylene-Di-Cl | C12Cl2H6 | 219.9847 |
| Acenaphthylene-Tri-Cl | C12Cl3H5 | 253.9457 |
| Acenaphthylene-Tetra-Cl | C12Cl4H4 | 287.9067 |
| Acenaphthylene-Mono-Br | C12BrH7 | 229.9731 |
| Acenaphthylene-Di-Br | C12Br2H6 | 307.8836 |
| Acenaphthylene-Tri-Br | C12Br3H5 | 385.7941 |
| Acenaphthylene-Tetra-Br | C12Br4H4 | 463.7047 |
| Acenaphthylene-Mono-Cl-Mono-Br | C12BrClH6 | 263.9341 |
| Acenaphthylene-Di-Cl-Mono-Br | C12BrCl2H5 | 297.8952 |
| Acenaphthylene-Mono-Cl-Di-Br | C12Br2ClH5 | 341.8447 |
| Acenaphthylene-Di-Cl-Di-Br | C12Br2Cl2H4 | 375.8057 |
| Acenaphthene-Mono-Cl | C12ClH9 | 188.0393 |
| Acenaphthene-Di-Cl | C12Cl2H8 | 222.0003 |
| Acenaphthene-Tri-Cl | C12Cl3H7 | 255.9613 |
| Acenaphthene-Tetra-Cl | C12Cl4H6 | 289.9224 |
| Acenaphthene-Mono-Br | C12BrH9 | 231.9888 |
| Acenaphthene-Di-Br | C12Br2H8 | 309.8993 |
| Acenaphthene-Tri-Br | C12Br3H7 | 387.8098 |
| Acenaphthene-Tetra-Br | C12Br4H6 | 465.7203 |
| Acenaphthene-Mono-Cl-Mono-Br | C12BrClH8 | 265.9498 |
| Acenaphthene-Di-Cl-Mono-Br | C12BrCl2H7 | 299.9108 |
| Acenaphthene-Mono-Cl-Di-Br | C12Br2ClH7 | 343.8603 |
| Acenaphthene-Di-Cl-Di-Br | C12Br2Cl2H6 | 377.8213 |

| | | |
|---|---|---|
| Fluorene-Mono-Cl | C13ClH9 | 200.0393 |
| Fluorene-Di-Cl | C13Cl2H8 | 234.0003 |
| Fluorene-Tri-Cl | C13Cl3H7 | 267.9613 |
| Fluorene-Tetra-Cl | C13Cl4H6 | 301.9224 |
| Fluorene-Mono-Br | C13BrH9 | 243.9888 |
| Fluorene-Di-Br | C13Br2H8 | 321.8993 |
| Fluorene-Tri-Br | C13Br3H7 | 399.8098 |
| Fluorene-Tetra-Br | C13Br4H6 | 477.7203 |
| Fluorene-Mono-Cl-Mono-Br | C13BrClH8 | 277.9498 |
| Fluorene-Di-Cl-Mono-Br | C13BrCl2H7 | 311.9108 |
| Fluorene-Mono-Cl-Di-Br | C13Br2ClH7 | 355.8603 |
| Fluorene-Di-Cl-Di-Br | C13Br2Cl2H6 | 389.8213 |
| Phenanthrene-Mono-Cl | C14ClH9 | 212.0393 |
| Phenanthrene-Di-Cl | C14Cl2H8 | 246.0003 |
| Phenanthrene-Tri-Cl | C14Cl3H7 | 279.9613 |
| Phenanthrene-Tetra-Cl | C14Cl4H6 | 313.9224 |
| Phenanthrene-Mono-Br | C14BrH9 | 255.9888 |
| Phenanthrene-Di-Br | C14Br2H8 | 333.8993 |
| Phenanthrene-Tri-Br | C14Br3H7 | 411.8098 |
| Phenanthrene-Tetra-Br | C14Br4H6 | 489.7203 |
| Phenanthrene-Mono-Cl-Mono-Br | C14BrClH8 | 289.9498 |
| Phenanthrene-Di-Cl-Mono-Br | C14BrCl2H7 | 323.9108 |
| Phenanthrene-Mono-Cl-Di-Br | C14Br2ClH7 | 367.8603 |
| Phenanthrene-Di-Cl-Di-Br | C14Br2Cl2H6 | 401.8213 |
| Fluoranthene-Mono-Cl | C16ClH9 | 236.0393 |
| Fluoranthene-Di-Cl | C16Cl2H8 | 270.0003 |
| Fluoranthene-Tri-Cl | C16Cl3H7 | 303.9613 |
| Fluoranthene-Tetra-Cl | C16Cl4H6 | 337.9224 |
| Fluoranthene-Mono-Br | C16BrH9 | 279.9888 |
| Fluoranthene-Di-Br | C16Br2H8 | 357.8993 |
| Fluoranthene-Tri-Br | C16Br3H7 | 435.8098 |
| Fluoranthene-Tetra-Br | C16Br4H6 | 513.7203 |
| Fluoranthene-Mono-Cl-Mono-Br | C16BrClH8 | 313.9498 |
| Fluoranthene-Di-Cl-Mono-Br | C16BrCl2H7 | 347.9108 |
| Fluoranthene-Mono-Cl-Di-Br | C16Br2ClH7 | 391.8603 |
| Fluoranthene-Di-Cl-Di-Br | C16Br2Cl2H6 | 425.8213 |
| Chrysene-Mono-Cl | C18ClH11 | 262.0549 |
| Chrysene-Di-Cl | C18Cl2H10 | 296.016 |
| Chrysene-Tri-Cl | C18Cl3H9 | 329.977 |
| Chrysene-Tetra-Cl | C18Cl4H8 | 363.938 |
| Chrysene-Mono-Br | C18BrH11 | 306.0044 |
| Chrysene-Di-Br | C18Br2H10 | 383.9149 |
| Chrysene-Tri-Br | C18Br3H9 | 461.8254 |
| Chrysene-Tetra-Br | C18Br4H8 | 539.736 |
| Chrysene-Mono-Cl-Mono-Br | C18BrClH10 | 339.9654 |
| Chrysene-Di-Cl-Mono-Br | C18BrCl2H9 | 373.9265 |
| Chrysene-Mono-Cl-Di-Br | C18Br2ClH9 | 417.876 |
| Chrysene-Di-Cl-Di-Br | C18Br2Cl2H8 | 451.837 |
| Benzo[a]pyrene-Mono-Cl | C20ClH11 | 286.0549 |
| Benzo[a]pyrene-Di-Cl | C20Cl2H10 | 320.016 |
| Benzo[a]pyrene-Tri-Cl | C20Cl3H9 | 353.977 |
| Benzo[a]pyrene-Tetra-Cl | C20Cl4H8 | 387.938 |
| Benzo[a]pyrene-Mono-Br | C20BrH11 | 330.0044 |
| Benzo[a]pyrene-Di-Br | C20Br2H10 | 407.9149 |
| Benzo[a]pyrene-Tri-Br | C20Br3H9 | 485.8254 |
| Benzo[a]pyrene-Tetra-Br | C20Br4H8 | 563.736 |
| Benzo[a]pyrene-Mono-Cl-Mono-Br | C20BrClH10 | 363.9654 |
| Benzo[a]pyrene-Di-Cl-Mono-Br | C20BrCl2H9 | 397.9265 |
| Benzo[a]pyrene-Mono-Cl-Di-Br | C20Br2ClH9 | 441.876 |
| Benzo[a]pyrene-Di-Cl-Di-Br | C20Br2Cl2H8 | 475.837 |
| Indeno[1,2,3-cd]pyrene-Mono-Cl | C22ClH11 | 310.0549 |
| Indeno[1,2,3-cd]pyrene-Di-Cl | C22Cl2H10 | 344.016 |

| | | |
|---|---|---|
| Indeno[1,2,3-cd]pyrene-Tri-Cl | C22Cl3H9 | 377.977 |
| Indeno[1,2,3-cd]pyrene-Tetra-Cl | C22Cl4H8 | 411.938 |
| Indeno[1,2,3-cd]pyrene-Mono-Br | C22BrH11 | 354.0044 |
| Indeno[1,2,3-cd]pyrene-Di-Br | C22Br2H10 | 431.9149 |
| Indeno[1,2,3-cd]pyrene-Tri-Br | C22Br3H9 | 509.8254 |
| Indeno[1,2,3-cd]pyrene-Tetra-Br | C22Br4H8 | 587.736 |
| Indeno[1,2,3-cd]pyrene-Mono-Cl-Mono-Br | C22BrClH10 | 387.9654 |
| Indeno[1,2,3-cd]pyrene-Di-Cl-Mono-Br | C22BrCl2H9 | 421.9265 |
| Indeno[1,2,3-cd]pyrene-Mono-Cl-Di-Br | C22Br2ClH9 | 465.876 |
| Indeno[1,2,3-cd]pyrene-Di-Cl-Di-Br | C22Br2Cl2H8 | 499.837 |
| Dibenz[a,h]anthracene-Mono-Cl | C22ClH13 | 312.0706 |
| Dibenz[a,h]anthracene-Di-Cl | C22Cl2H12 | 346.0316 |
| Dibenz[a,h]anthracene-Tri-Cl | C22Cl3H11 | 379.9926 |
| Dibenz[a,h]anthracene-Tetra-Cl | C22Cl4H10 | 413.9537 |
| Dibenz[a,h]anthracene-Mono-Br | C22BrH13 | 356.0201 |
| Dibenz[a,h]anthracene-Di-Br | C22Br2H12 | 433.9306 |
| Dibenz[a,h]anthracene-Tri-Br | C22Br3H11 | 511.8411 |
| Dibenz[a,h]anthracene-Tetra-Br | C22Br4H10 | 589.7516 |
| Dibenz[a,h]anthracene-Mono-Cl-Mono-Br | C22BrClH12 | 389.9811 |
| Dibenz[a,h]anthracene-Di-Cl-Mono-Br | C22BrCl2H11 | 423.9421 |
| Dibenz[a,h]anthracene-Mono-Cl-Di-Br | C22Br2ClH11 | 467.8916 |
| Dibenz[a,h]anthracene-Di-Cl-Di-Br | C22Br2Cl2H10 | 501.8526 |
| Benzo[g,h,i]perylene-Mono-Cl | C22ClH11 | 310.0549 |
| Benzo[g,h,i]perylene-Di-Cl | C22Cl2H10 | 344.016 |
| Benzo[g,h,i]perylene-Tri-Cl | C22Cl3H9 | 377.977 |
| Benzo[g,h,i]perylene-Tetra-Cl | C22Cl4H8 | 411.938 |
| Benzo[g,h,i]perylene-Mono-Br | C22BrH11 | 354.0044 |
| Benzo[g,h,i]perylene-Di-Br | C22Br2H10 | 431.9149 |
| Benzo[g,h,i]perylene-Tri-Br | C22Br3H9 | 509.8254 |
| Benzo[g,h,i]perylene-Tetra-Br | C22Br4H8 | 587.736 |
| Benzo[g,h,i]perylene-Mono-Cl-Mono-Br | C22BrClH10 | 387.9654 |
| Benzo[g,h,i]perylene-Di-Cl-Mono-Br | C22BrCl2H9 | 421.9265 |
| Benzo[g,h,i]perylene-Mono-Cl-Di-Br | C22Br2ClH9 | 465.876 |
| Benzo[g,h,i]perylene-Di-Cl-Di-Br | C22Br2Cl2H8 | 499.837 |
| Indole-Mono-Cl | C8ClH6N | 151.0189 |
| Indole-Di-Cl | C8Cl2H5N | 184.9799 |
| Indole-Tri-Cl | C8Cl3H4N | 218.9409 |
| Indole-Tetra-Cl | C8Cl4H3N | 252.902 |
| Indole-Mono-Br | C8BrH6N | 194.9684 |
| Indole-Di-Br | C8Br2H5N | 272.8789 |
| Indole-Tri-Br | C8Br3H4N | 350.7894 |
| Indole-Tetra-Br | C8Br4H3N | 428.6999 |
| Indole-Mono-Cl-Mono-Br | C8BrClH5N | 228.9294 |
| Indole-Di-Cl-Mono-Br | C8BrCl2H4N | 262.8904 |
| Indole-Mono-Cl-Di-Br | C8Br2ClH4N | 306.8399 |
| Indole-Di-Cl-Di-Br | C8Br2Cl2H3N | 340.8009 |
| Quinoline-Mono-Cl | C9ClH6N | 163.0189 |
| Quinoline-Di-Cl | C9Cl2H5N | 196.9799 |
| Quinoline-Tri-Cl | C9Cl3H4N | 230.9409 |
| Quinoline-Tetra-Cl | C9Cl4H3N | 264.902 |
| Quinoline-Mono-Br | C9BrH6N | 206.9684 |
| Quinoline-Di-Br | C9Br2H5N | 284.8789 |
| Quinoline-Tri-Br | C9Br3H4N | 362.7894 |
| Quinoline-Tetra-Br | C9Br4H3N | 440.6999 |
| Quinoline-Mono-Cl-Mono-Br | C9BrClH5N | 240.9294 |
| Quinoline-Di-Cl-Mono-Br | C9BrCl2H4N | 274.8904 |
| Quinoline-Mono-Cl-Di-Br | C9Br2ClH4N | 318.8399 |
| Quinoline-Di-Cl-Di-Br | C9Br2Cl2H3N | 352.8009 |
| Acridine-Mono-Cl | C13ClH8N | 213.0345 |
| Acridine-Di-Cl | C13Cl2H7N | 246.9956 |
| Acridine-Tri-Cl | C13Cl3H6N | 280.9566 |
| Acridine-Tetra-Cl | C13Cl4H5N | 314.9176 |

| | | |
|---|---|---|
| Acridine-Mono-Br | C13BrH8N | 256.984 |
| Acridine-Di-Br | C13Br2H7N | 334.8945 |
| Acridine-Tri-Br | C13Br3H6N | 412.805 |
| Acridine-Tetra-Br | C13Br4H5N | 490.7156 |
| Acridine-Mono-Cl-Mono-Br | C13BrClH7N | 290.945 |
| Acridine-Di-Cl-Mono-Br | C13BrCl2H6N | 324.9061 |
| Acridine-Mono-Cl-Di-Br | C13Br2ClH6N | 368.8556 |
| Acridine-Di-Cl-Di-Br | C13Br2Cl2H5N | 402.8166 |
| Carbazole-Mono-Cl | C12ClH8N | 201.0345 |
| Carbazole-Di-Cl | C12Cl2H7N | 234.9956 |
| Carbazole-Tri-Cl | C12Cl3H6N | 268.9566 |
| Carbazole-Tetra-Cl | C12Cl4H5N | 302.9176 |
| Carbazole-Mono-Br | C12BrH8N | 244.984 |
| Carbazole-Di-Br | C12Br2H7N | 322.8945 |
| Carbazole-Tri-Br | C12Br3H6N | 400.805 |
| Carbazole-Tetra-Br | C12Br4H5N | 478.7156 |
| Carbazole-Mono-Cl-Mono-Br | C12BrClH7N | 278.945 |
| Carbazole-Di-Cl-Mono-Br | C12BrCl2H6N | 312.9061 |
| Carbazole-Mono-Cl-Di-Br | C12Br2ClH6N | 356.8556 |
| Carbazole-Di-Cl-Di-Br | C12Br2Cl2H5N | 390.8166 |
| Benz(a)acridine-Mono-Cl | C17ClH10N | 263.0502 |
| Benz(a)acridine-Di-Cl | C17Cl2H9N | 297.0112 |
| Benz(a)acridine-Tri-Cl | C17Cl3H8N | 330.9722 |
| Benz(a)acridine-Tetra-Cl | C17Cl4H7N | 364.9333 |
| Benz(a)acridine-Mono-Br | C17BrH10N | 306.9997 |
| Benz(a)acridine-Di-Br | C17Br2H9N | 384.9102 |
| Benz(a)acridine-Tri-Br | C17Br3H8N | 462.8207 |
| Benz(a)acridine-Tetra-Br | C17Br4H7N | 540.7312 |
| Benz(a)acridine-Mono-Cl-Mono-Br | C17BrClH9N | 340.9607 |
| Benz(a)acridine-Di-Cl-Mono-Br | C17BrCl2H8N | 374.9217 |
| Benz(a)acridine-Mono-Cl-Di-Br | C17Br2ClH8N | 418.8712 |
| Benz(a)acridine-Di-Cl-Di-Br | C17Br2Cl2H7N | 452.8322 |
| 11H-Benzo[a]carbazole-Mono-Cl | C16ClH10N | 251.0502 |
| 11H-Benzo[a]carbazole-Di-Cl | C16Cl2H9N | 285.0112 |
| 11H-Benzo[a]carbazole-Tri-Cl | C16Cl3H8N | 318.9722 |
| 11H-Benzo[a]carbazole-Tetra-Cl | C16Cl4H7N | 352.9333 |
| 11H-Benzo[a]carbazole-Mono-Br | C16BrH10N | 294.9997 |
| 11H-Benzo[a]carbazole-Di-Br | C16Br2H9N | 372.9102 |
| 11H-Benzo[a]carbazole-Tri-Br | C16Br3H8N | 450.8207 |
| 11H-Benzo[a]carbazole-Tetra-Br | C16Br4H7N | 528.7312 |
| 11H-Benzo[a]carbazole-Mono-Cl-Mono-Br | C16BrClH9N | 328.9607 |
| 11H-Benzo[a]carbazole-Di-Cl-Mono-Br | C16BrCl2H8N | 362.9217 |
| 11H-Benzo[a]carbazole-Mono-Cl-Di-Br | C16Br2ClH8N | 406.8712 |
| 11H-Benzo[a]carbazole-Di-Cl-Di-Br | C16Br2Cl2H7N | 440.8322 |
| Dibenz[a,j]acridine-Mono-Cl | C21ClH12N | 313.0658 |
| Dibenz[a,j]acridine-Di-Cl | C21Cl2H11N | 347.0269 |
| Dibenz[a,j]acridine-Tri-Cl | C21Cl3H10N | 380.9879 |
| Dibenz[a,j]acridine-Tetra-Cl | C21Cl4H9N | 414.9489 |
| Dibenz[a,j]acridine-Mono-Br | C21BrH12N | 357.0153 |
| Dibenz[a,j]acridine-Di-Br | C21Br2H11N | 434.9258 |
| Dibenz[a,j]acridine-Tri-Br | C21Br3H10N | 512.8363 |
| Dibenz[a,j]acridine-Tetra-Br | C21Br4H9N | 590.7469 |
| Dibenz[a,j]acridine-Mono-Cl-Mono-Br | C21BrClH11N | 390.9763 |
| Dibenz[a,j]acridine-Di-Cl-Mono-Br | C21BrCl2H10N | 424.9374 |
| Dibenz[a,j]acridine-Mono-Cl-Di-Br | C21Br2ClH10N | 468.8869 |
| Dibenz[a,j]acridine-Di-Cl-Di-Br | C21Br2Cl2H9N | 502.8479 |
| 7H-Dibenzo[c,g]carbazole-Mono-Cl | C20ClH12N | 301.0658 |
| 7H-Dibenzo[c,g]carbazole-Di-Cl | C20Cl2H11N | 335.0269 |
| 7H-Dibenzo[c,g]carbazole-Tri-Cl | C20Cl3H10N | 368.9879 |
| 7H-Dibenzo[c,g]carbazole-Tetra-Cl | C20Cl4H9N | 402.9489 |
| 7H-Dibenzo[c,g]carbazole-Mono-Br | C20BrH12N | 345.0153 |
| 7H-Dibenzo[c,g]carbazole-Di-Br | C20Br2H11N | 422.9258 |

| | | |
|---|---|---|
| 7H-Dibenzo[c,g]carbazole-Tri-Br | C20Br3H10N | 500.8363 |
| 7H-Dibenzo[c,g]carbazole-Tetra-Br | C20Br4H9N | 578.7469 |
| 7H-Dibenzo[c,g]carbazole-Mono-Cl-Mono-Br | C20BrClH11N | 378.9763 |
| 7H-Dibenzo[c,g]carbazole-Di-Cl-Mono-Br | C20BrCl2H10N | 412.9374 |
| 7H-Dibenzo[c,g]carbazole-Mono-Cl-Di-Br | C20Br2ClH10N | 456.8869 |
| 7H-Dibenzo[c,g]carbazole-Di-Cl-Di-Br | C20Br2Cl2H9N | 490.8479 |
| Benzothiophene-Mono-Cl | C8ClH5S | 167.9801 |
| Benzothiophene-Di-Cl | C8Cl2H4S | 201.9411 |
| Benzothiophene-Tri-Cl | C8Cl3H3S | 235.9021 |
| Benzothiophene-Tetra-Cl | C8Cl4H2S | 269.8631 |
| Benzothiophene-Mono-Br | C8BrH5S | 211.9295 |
| Benzothiophene-Di-Br | C8Br2H4S | 289.8401 |
| Benzothiophene-Tri-Br | C8Br3H3S | 367.7506 |
| Benzothiophene-Tetra-Br | C8Br4H2S | 445.6611 |
| Benzothiophene-Mono-Cl-Mono-Br | C8BrClH4S | 245.8906 |
| Benzothiophene-Di-Cl-Mono-Br | C8BrCl2H3S | 279.8516 |
| Benzothiophene-Mono-Cl-Di-Br | C8Br2ClH3S | 323.8011 |
| Benzothiophene-Di-Cl-Di-Br | C8Br2Cl2H2S | 357.7621 |
| Dibenzothiophene-Mono-Cl | C12ClH7S | 217.9957 |
| Dibenzothiophene-Di-Cl | C12Cl2H6S | 251.9567 |
| Dibenzothiophene-Tri-Cl | C12Cl3H5S | 285.9178 |
| Dibenzothiophene-Tetra-Cl | C12Cl4H4S | 319.8788 |
| Dibenzothiophene-Mono-Br | C12BrH7S | 261.9452 |
| Dibenzothiophene-Di-Br | C12Br2H6S | 339.8557 |
| Dibenzothiophene-Tri-Br | C12Br3H5S | 417.7662 |
| Dibenzothiophene-Tetra-Br | C12Br4H4S | 495.6767 |
| Dibenzothiophene-Mono-Cl-Mono-Br | C12BrClH6S | 295.9062 |
| Dibenzothiophene-Di-Cl-Mono-Br | C12BrCl2H5S | 329.8672 |
| Dibenzothiophene-Mono-Cl-Di-Br | C12Br2ClH5S | 373.8167 |
| Dibenzothiophene-Di-Cl-Di-Br | C12Br2Cl2H4S | 407.7778 |
| Benzo(b)naphtho(1,2-d)thiophene-Mono-Cl | C16ClH9S | 268.0114 |
| Benzo(b)naphtho(1,2-d)thiophene-Di-Cl | C16Cl2H8S | 301.9724 |
| Benzo(b)naphtho(1,2-d)thiophene-Tri-Cl | C16Cl3H7S | 335.9334 |
| Benzo(b)naphtho(1,2-d)thiophene-Tetra-Cl | C16Cl4H6S | 369.8944 |
| Benzo(b)naphtho(1,2-d)thiophene-Mono-Br | C16BrH9S | 311.9608 |
| Benzo(b)naphtho(1,2-d)thiophene-Di-Br | C16Br2H8S | 389.8714 |
| Benzo(b)naphtho(1,2-d)thiophene-Tri-Br | C16Br3H7S | 467.7819 |
| Benzo(b)naphtho(1,2-d)thiophene-Tetra-Br | C16Br4H6S | 545.6924 |
| Benzo(b)naphtho(1,2-d)thiophene-Mono-Cl-Mono-Br | C16BrClH8S | 345.9219 |
| Benzo(b)naphtho(1,2-d)thiophene-Di-Cl-Mono-Br | C16BrCl2H7S | 379.8829 |
| Benzo(b)naphtho(1,2-d)thiophene-Mono-Cl-Di-Br | C16Br2ClH7S | 423.8324 |
| Benzo(b)naphtho(1,2-d)thiophene-Di-Cl-Di-Br | C16Br2Cl2H6S | 457.7934 |
| Benzofuran-Mono-Cl | C8ClH5O | 152.0029 |
| Benzofuran-Di-Cl | C8Cl2H4O | 185.9639 |
| Benzofuran-Tri-Cl | C8Cl3H3O | 219.925 |
| Benzofuran-Tetra-Cl | C8Cl4H2O | 253.886 |
| Benzofuran-Mono-Br | C8BrH5O | 195.9524 |
| Benzofuran-Di-Br | C8Br2H4O | 273.8629 |
| Benzofuran-Tri-Br | C8Br3H3O | 351.7734 |
| Benzofuran-Tetra-Br | C8Br4H2O | 429.6839 |
| Benzofuran-Mono-Cl-Mono-Br | C8BrClH4O | 229.9134 |
| Benzofuran-Di-Cl-Mono-Br | C8BrCl2H3O | 263.8744 |
| Benzofuran-Mono-Cl-Di-Br | C8Br2ClH3O | 307.8239 |
| Benzofuran-Di-Cl-Di-Br | C8Br2Cl2H2O | 341.785 |
| Dibenzofuran-Mono-Cl | C12ClH7O | 202.0185 |
| Dibenzofuran-Di-Cl | C12Cl2H6O | 235.9796 |
| Dibenzofuran-Tri-Cl | C12Cl3H5O | 269.9406 |
| Dibenzofuran-Tetra-Cl | C12Cl4H4O | 303.9016 |
| Dibenzofuran-Mono-Br | C12BrH7O | 245.968 |
| Dibenzofuran-Di-Br | C12Br2H6O | 323.8785 |
| Dibenzofuran-Tri-Br | C12Br3H5O | 401.7891 |
| Dibenzofuran-Tetra-Br | C12Br4H4O | 479.6996 |

| | | |
|---|---|---|
| Dibenzofuran-Mono-Cl-Mono-Br | C12BrClH6O | 279.9291 |
| Dibenzofuran-Di-Cl-Mono-Br | C12BrCl2H5O | 313.8901 |
| Dibenzofuran-Mono-Cl-Di-Br | C12Br2ClH5O | 357.8396 |
| Dibenzofuran-Di-Cl-Di-Br | C12Br2Cl2H4O | 391.8006 |
| Benzonaphthodibenzofuran-Mono-Cl | C16ClH9O | 252.0342 |
| Benzonaphthodibenzofuran-Di-Cl | C16Cl2H8O | 285.9952 |
| Benzonaphthodibenzofuran-Tri-Cl | C16Cl3H7O | 319.9563 |
| Benzonaphthodibenzofuran-Tetra-Cl | C16Cl4H6O | 353.9173 |
| Benzonaphthodibenzofuran-Mono-Br | C16BrH9O | 295.9837 |
| Benzonaphthodibenzofuran-Di-Br | C16Br2H8O | 373.8942 |
| Benzonaphthodibenzofuran-Tri-Br | C16Br3H7O | 451.8047 |
| Benzonaphthodibenzofuran-Tetra-Br | C16Br4H6O | 529.7152 |
| Benzonaphthodibenzofuran-Mono-Cl-Mono-Br | C16BrClH8O | 329.9447 |
| Benzonaphthodibenzofuran-Di-Cl-Mono-Br | C16BrCl2H7O | 363.9057 |
| Benzonaphthodibenzofuran-Mono-Cl-Di-Br | C16Br2ClH7O | 407.8552 |
| Benzonaphthodibenzofuran-Di-Cl-Di-Br | C16Br2Cl2H6O | 441.8163 |

**Methodology of Calculating Performance Characteristics of Workflow Benchmarking Tests**

Four metrics are reported throughout this work: sensitivity (recall), specificity, precision, and the $F_1$ score. All four are computed from a standard 2×2 confusion matrix, defined by true positives (TP), false negatives (FN), false positives (FP), and true negatives (TN):

$$Sensitivity = TP / (TP + FN)$$

$$Specificity = TN / (TN + FP)$$

$$Precision = TP / (TP + FP)$$

$$F_1 = 2 \cdot Precision \cdot Sensitivity / (Precision + Sensitivity)$$

What differs between the six benchmarks below is not the arithmetic but the definition of the population, the positive class, and what counts as "flagged." Each section states these definitions explicitly before presenting the confusion matrix. Benchmark sections use a full confusion matrix (a defined positive and negative class exists in the reference data).

Benchmark Section 1 - Aromaticity Index Filter

This benchmark evaluates the aromaticity index (AI) filter in the specific decision it makes within WF2: discriminating halogenated aromatic compounds from halogenated aliphatic compounds among candidates that have already passed the halogen isotope-pattern filter. The population is therefore restricted to the 96 halogenated compounds in the 505-compound in silico reference set; non-halogenated compounds are outside the scope of this test because they are removed upstream by isotope screening and never reach the AI filter.

Positive class: halogenated aromatic (n = 70) — compounds the filter should retain.

Negative class: halogenated aliphatic (n = 26) — compounds the filter should reject.

Flagged: compounds with $-0.05 < AI < 0.005$ (the WF2 acceptance window).

Among the 96 halogenated compounds evaluated, the aromaticity index (AI) filter flagged 83 compounds as being within the WF2 acceptance window ($-0.05 < AI < 0.005$). Of these, 58 were correctly identified as halogenated aromatic compounds (true positives), while 25 were incorrectly retained despite being halogenated aliphatic compounds (false positives). The remaining 13 compounds fell outside the acceptance window; this group comprised 12 halogenated aromatic compounds that were incorrectly rejected (false negatives) and one halogenated aliphatic compound that was correctly excluded (true negative). Overall, the reference set contained 70 halogenated aromatic compounds (positive class) and 26 halogenated aliphatic compounds (negative class).

$$Sensitivity = 58 / 70 = 0.829$$

$$Precision = 58 / 83 = 0.699$$

$$F_1 = 2(0.699)(0.829) / (0.699+0.829) = 0.758$$

Specificity is not reported in the main text for this benchmark. Computed directly, specificity = TN/(TN+FP) = 1/26 = 0.038: the AI filter alone excludes only 1 of 26 halogenated aliphatics. This number was omitted as a headline statistic because, presented without context, it reads as a failure of the filter. The substantive finding is that the AI cannot by itself separate halogenated aliphatics from halogenated aromatics. That this is precisely why WF2 does not rely on the AI in isolation and is instead stated explicitly in the main text rather than left to an isolated number to convey.

Benchmark Section 2 - Empirical Validation of WF2 Isotope Pattern Filter with RECETOX Library

This benchmark evaluates whether WF2's isotope-pattern filter correctly distinguishes halogenated from non-halogenated compounds, using the full 386 compound RECETOX Exposome library as an independent, externally sourced test set.

The isotope-pattern filter identified 209 compounds as halogen-bearing compounds. Of these, 178 were correctly identified as halogenated compounds (true positives), while 31 were incorrectly flagged despite being non-halogenated (false positives). The remaining 177 compounds were not flagged by the filter, comprising 13 halogenated compounds that were missed (false negatives) and 164 non-halogenated compounds that were correctly rejected (true negatives). Thus, the reference library contained 191 halogenated compounds (positive class) and 195 non-halogenated compounds (negative class).

*Sensitivity = 178/191 = 0.932*

*Specificity = 164/195 = 0.841*

*Precision = 178/209 = 0.852*

$F_1 = 0.890$

Twelve of the 191 halogenated compounds contain fluorine as their only halogen. Because $^{19}F$ is monoisotopic, it produces no M+2 isotopologue and is therefore undetectable in principle by any isotope-pattern criterion. After excluding these fluorine-only compounds from the positive class, the adjusted reference set contained 179 halogenated compounds. Under this adjusted evaluation, the filter correctly identified 171 halogenated compounds (true positives) and missed 8 (false negatives), while the numbers of false positives (31) and true negatives (164) remained unchanged because the negative class was unaffected.

*Sensitivity = 171/179 = 0.955*

*Specificity = 164/195 = 0.841*

*Precision = 171/202 = 0.847*

$F_1 = 0.898$

Benchmark Section 3 - Empirical Validation of WF2 Exact Mass Matching with RECETOX Library

This benchmark evaluates the final WF2 output (after exact-mass matching against the halo-PAC reference table) against the narrower target class of fused halogenated PACs.

Positive class: fused halogenated PACs (n = 13, as defined by ring-fusion analysis in RDKit). Negative class: all other compounds in the library (n = 373). Flagged: the 96 candidates in the final halo-PAC output. Of these, 7 were correctly identified as fused halogenated PACs (true positives), whereas 89 were incorrectly flagged despite belonging to other chemical classes (false positives). The remaining 290 compounds were not flagged by the workflow, comprising 6 fused halogenated PACs that were missed (false negatives) and 284 compounds that were correctly excluded (true negatives). Overall, the reference library contained 13 fused halogenated PACs (positive class) and 373 compounds that did not belong to the target class (negative class).

*Sensitivity = 7/13 = 0.538*

*Specificity = 284/373 = 0.761*

*Precision = 7/96 = 0.073*

Benchmark Section 4 - Empirical Validation of WF3 with RECETOX Library

This benchmark evaluates the heterocyclic PAC (HPAC) screening logic of WF3 at four evidence-score thresholds. Positive class: fused heterocyclic PACs (n = 23, ring N/S/O atom required). Negative class: all other compounds (n = 363). Base rate = 23/386 = 0.0596.

The performance of WF3 was evaluated at four evidence-score thresholds (≥ 3, ≥ 5, ≥ 6, and ≥ 7) using the 386-compound RECETOX Exposome library. The reference set contained 23 fused heterocyclic PACs (positive class) and 363 compounds that did not belong to the target class (negative class), corresponding to a base rate of 23/386 (0.0596).

At an evidence-score threshold of ≥ 3, WF3 flagged 354 compounds as candidates. This included 23 true positives and 331 false positives, while no heterocyclic PACs were missed (0 false negatives) and 32 compounds were correctly rejected (true negatives).

At the reported threshold of ≥ 5, WF3 flagged 293 compounds. Of these, 21 were correctly identified as fused heterocyclic PACs (true positives) and 272 were false positives. Two heterocyclic PACs were missed (false negatives), while 91 non-target compounds were correctly excluded (true negatives). This threshold gave a sensitivity of 0.913, a specificity of 0.251, a precision of 0.072, an $F_1$ score of 0.133, and an enrichment factor of 1.20 relative to the base rate.

At a threshold of ≥ 6, the number of flagged candidates decreased to 128. This consisted of 11 true positives and 117 false positives, while 12 heterocyclic PACs were missed and 246 non-target compounds were correctly rejected.

At the most stringent threshold of ≥ 7, only 12 compounds were flagged. These included one true positive and 11 false positives, while 22 heterocyclic PACs were missed and 352 non-target compounds were correctly excluded.

These results illustrate the expected trade-off between sensitivity and specificity. Lower evidence-score thresholds maximise the recovery of heterocyclic PACs but generate many false-positive candidates, whereas higher thresholds substantially reduce the number of candidates at the cost of missing an increasing proportion of the target compounds. The

threshold of ≥ 5 was selected as the operating point because it provides a balance between high sensitivity and manageable candidate numbers while maintaining the intended first-pass screening role of WF3

*Worked example at threshold ≥ 5:*

*Sensitivity = 21/23 = 0.913;*

*Specificity = 91/363 = 0.251*

*Precision = 21/293 = 0.072;*

*Enrichment = Precision / Base rate = 0.072 / 0.0596 = 1.20×*

Enrichment (precision divided by the base rate of the target class) is introduced specifically for WF3 because precision alone is difficult to interpret at a 6% base rate: a precision of 0.072 sounds poor in isolation but represents only a modest 1.20× improvement over random selection at the reported threshold, and this framing communicates the filter's actual discriminating value more transparently than precision alone. All four rows use the identical confusion-matrix definitions; only the evidence-score cutoff used to define “flagged” changes.

**Table V 8:** Summary of the output file of the standard run checked through the workflow 2

| Halo-PAC standards | Aromaticity Index | Possible halogen class/ classes | Halo PAC matched | Matched formula | Mass accuracy (ppm) | Detected Isotope peaks |
|---|---|---|---|---|---|---|
| 5-Cl-BT | -0.002 | mono_Cl | Benzothiophene-Mono-Cl | C8ClH5S | 8.5 | 167.97863 (ri=1000), 169.97557 (ri=367) |
| 2-Cl-Qui | 0.001 | mono_Cl | Quinoline-Mono-Cl | C9ClH6N | 11.5 | 163.01700 (ri=660), 165.01416 (ri=218) |
| 2-Br-BT | -0.004 | mono_Br | Benzothiophene-Mono-Br | C8BrH5S | 7.1 | 211.92802 (ri=1000), 213.92594 (ri=941) |
| 2,3-Cl2-BT | -0.004 | mono_Cl; di_Cl; poly_or_mixed | Benzothiophene-Di-Cl | C8Cl2H4S | 2.9 | 201.94050 (ri=1000), 203.93750 (ri=614), 166.97186 (ri=411), 168.96873 (ri=148), 205.93449 (ri=118) |
| 2-Br-Qui | -0.002 | mono_Br | Quinoline-Mono-Br | C9BrH6N | 4.9 | 206.96734 (ri=215), 208.96521 (ri=211) |
| 5-Cl-Ind | 0.002 | mono_Cl | Indole-Mono-Cl | C8ClH6N | 8.8 | 151.01755 (ri=1000), 153.01465 (ri=318) |
| 3-Br-IQui | -0.002 | mono_Br | Quinoline-Mono-Br | C9BrH6N | 5.2 | 206.96729 (ri=405), 208.96530 (ri=335) |
| 5-Br-Ind | -0.002 | mono_Br | Indole-Mono-Br | C8BrH6N | 2.3 | 194.96791 (ri=911), 196.96570 (ri=895) |
| 5-Br-Ace | -0.001 | mono_Br | Acenaphthene-Mono-Br; Dibenzofuran-Di-Cl | C12BrH9; C12Cl2H6O | 6.6 | 231.98723 (ri=344), 233.98520 (ri=323) |
| 9-Br-Flu | -0.001 | mono_Br | Fluorene-Mono-Br | C13BrH9 | 10.9 | 243.98610 (ri=150), 245.98409 (ri=141) |
| 9-Cl-Acr | 0.002 | mono_Cl | Acridine-Mono-Cl | C13ClH8N | 12.2 | 213.03193 (ri=1000), 215.02910 (ri=347) |
| 1-Cl-Ant | 0.002 | mono_Cl | Phenanthrene-Mono-Cl | C14ClH9 | 10.9 | 212.03697 (ri=1000), 214.03409 (ri=309) |
| 2,7-Cl2-Flu | 0.000 | mono_Cl; poly_or_mixed | Fluorene-Di-Cl | C13Cl2H8 | 7.1 | 199.03010 (ri=1000), 201.02717 (ri=306), 233.99866 (ri=223), 235.99564 (ri=132) |
| 1,2-Br2-Acy | -0.005 | mono_Br; di_Br; mono_Cl | Acenaphthylene-Di-Br | C12Br2H6 | 5.0 | 309.88011 (ri=730), 307.88209 (ri=405), 311.87775 (ri=357), 230.96238 (ri=122), 228.96462 (ri=113) |
| 2-Br-DBT | -0.003 | mono_Br | Dibenzothiophene-Mono-Br | C12BrH7S | 2.8 | 261.94446 (ri=1000), 263.94220 (ri=959), 262.94750 (ri=147), 264.94528 (ri=131) |
| 9-Br-Acr | -0.001 | mono_Br | Phenanthrene-Mono-Br | C14BrH9 | 6.1 | 255.98721 (ri=1000), 257.98535 (ri=934), 256.99092 (ri=148), 258.98919 (ri=137) |
| 9-Br-Ant | -0.001 | mono_Br | Phenanthrene-Mono-Br | C14BrH9 | 2.5 | 255.98811 (ri=1000), 257.98598 (ri=975), 258.98900 (ri=162), 256.99131 (ri=158) |
| 1,4-Cl2-Ant | 0.000 | di_Cl; poly_or_mixed | Phenanthrene-Di-Cl | C14Cl2H8 | 18.0 | 245.99589 (ri=1000), 247.99319 (ri=585), 249.99013 (ri=104) |
| 5,6-Br2-1,2-H2-Acy | -0.004 | di_Br; mono_Cl | Acenaphthene-Di-Br; Dibenzofuran-Di-Cl-Mono-Br | C12Br2H8; C12BrCl2H5O | 4.5 | 311.89449 (ri=276), 309.89678 (ri=150), 313.89255 (ri=132) |
| 2,7-Br2-Flu | -0.004 | mono_Br; di_Br; mono_Cl | Fluorene-Di-Br | C13Br2H8 | 6.8 | 242.97898 (ri=1000), 244.97718 (ri=943), 323.89468 (ri=368), 321.89708 (ri=173), 325.89261 (ri=169), 243.98279 (ri=152), 245.98059 (ri=147) |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3,6-Cl2-Car | -0.001 | mono_Cl; di_Cl; poly_or_mixed | Carbazole-Di-Cl; Dibenzofuran-Mono-Cl | C12Cl2H7N; C12ClH7O | 12.0 | 234.99274 (ri=1000), 200.02437 (ri=740), 236.98998 (ri=616), 202.02135 (ri=247), 238.98702 (ri=110) |
| 9,10-Br2-Ant | -0.004 | di_Br; mono_Cl | Phenanthrene-Di-Br | C14Br2H8 | 8.7 | 335.89430 (ri=589), 333.89639 (ri=314), 337.89246 (ri=279) |
| 1-Br-Pyr | -0.001 | mono_Br | Fluoranthene-Mono-Br | C16BrH9 | 2.7 | 281.98592 (ri=768), 279.98801 (ri=764), 280.99130 (ri=149), 282.98928 (ri=125) |
| 3,6-Br2-Car | -0.004 | poly_or_mixed | Carbazole-Di-Br | C12Br2H7N | 18.6 | 324.88849 (ri=1000), 326.88734 (ri=517) |
| 9-Br-1,5-Cl-Ant | -0.004 | di_Br; mono_Cl | Phenanthrene-Di-Cl-Mono-Br | C14BrCl2H7 | 1.7 | 325.90725 (ri=211), 323.91027 (ri=125), 327.90385 (ri=102) |
| 1,5,9,10-Cl4-Ant | -0.003 | poly_or_mixed | Acridine-Tetra-Cl; Phenanthrene-Tetra-Cl | C13Cl4H5N; C14Cl4H6 | 10.0 | 315.91774 (ri=176), 317.91615 (ri=101) |
| 1,6-Br2-Pyr | -0.004 | mono_Cl | - | - | - | 359.89575 (ri=761), 361.89384 (ri=350) |
| 4-Br-BaA | 0.000 | mono_Br | Chrysene-Mono-Br | C18BrH11 | 5.9 | 308.00006 (ri=811), 306.00262 (ri=778) |

* ri = relative intensity

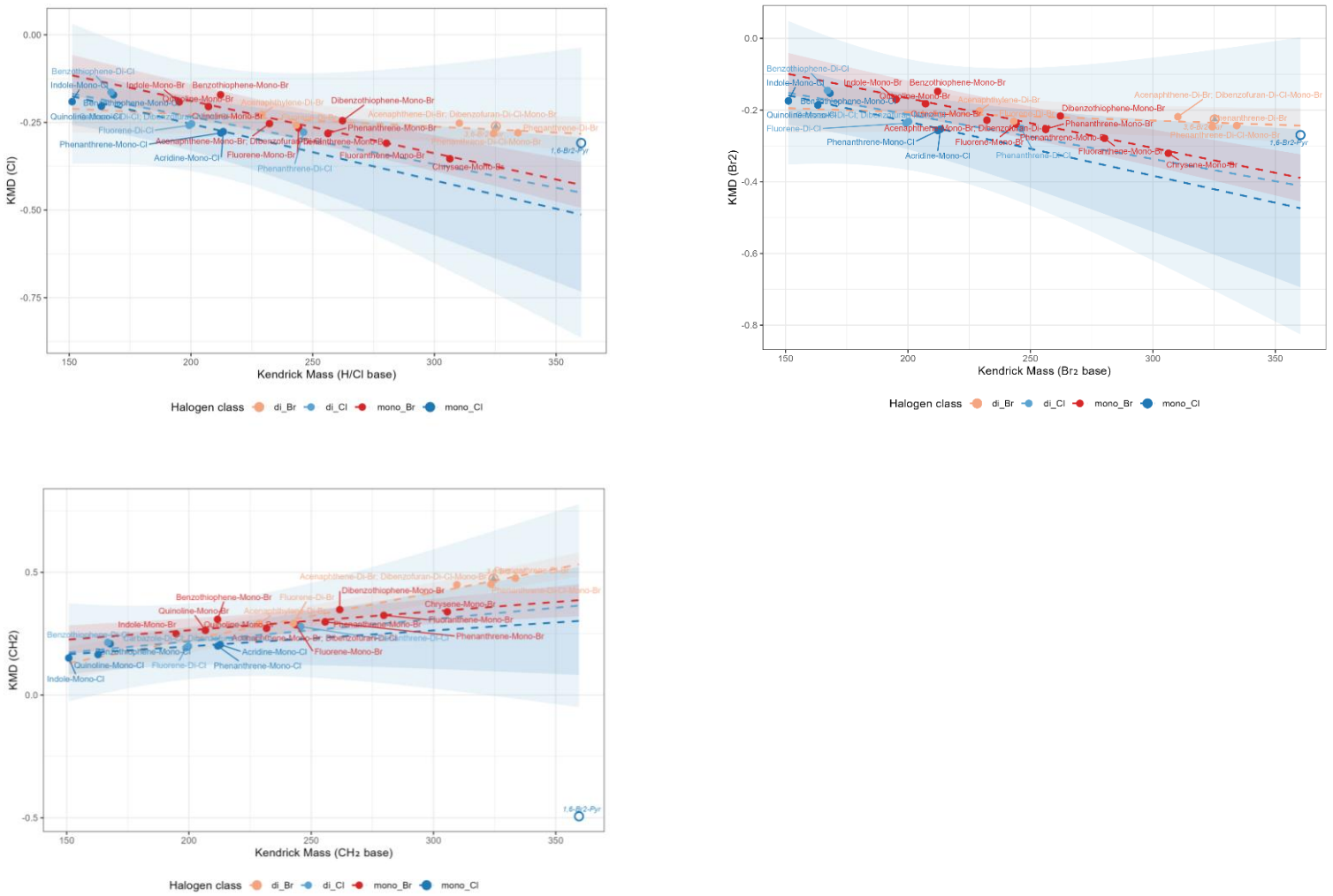



**Figure V 1:** Kendrick mass defect plots created for halogenated PACs standards using H/Cl (top panel), $Br_2$ (middle panel), and $CH_2$ (bottom panel). Colored bands show 95% predication intervals for each compound class.

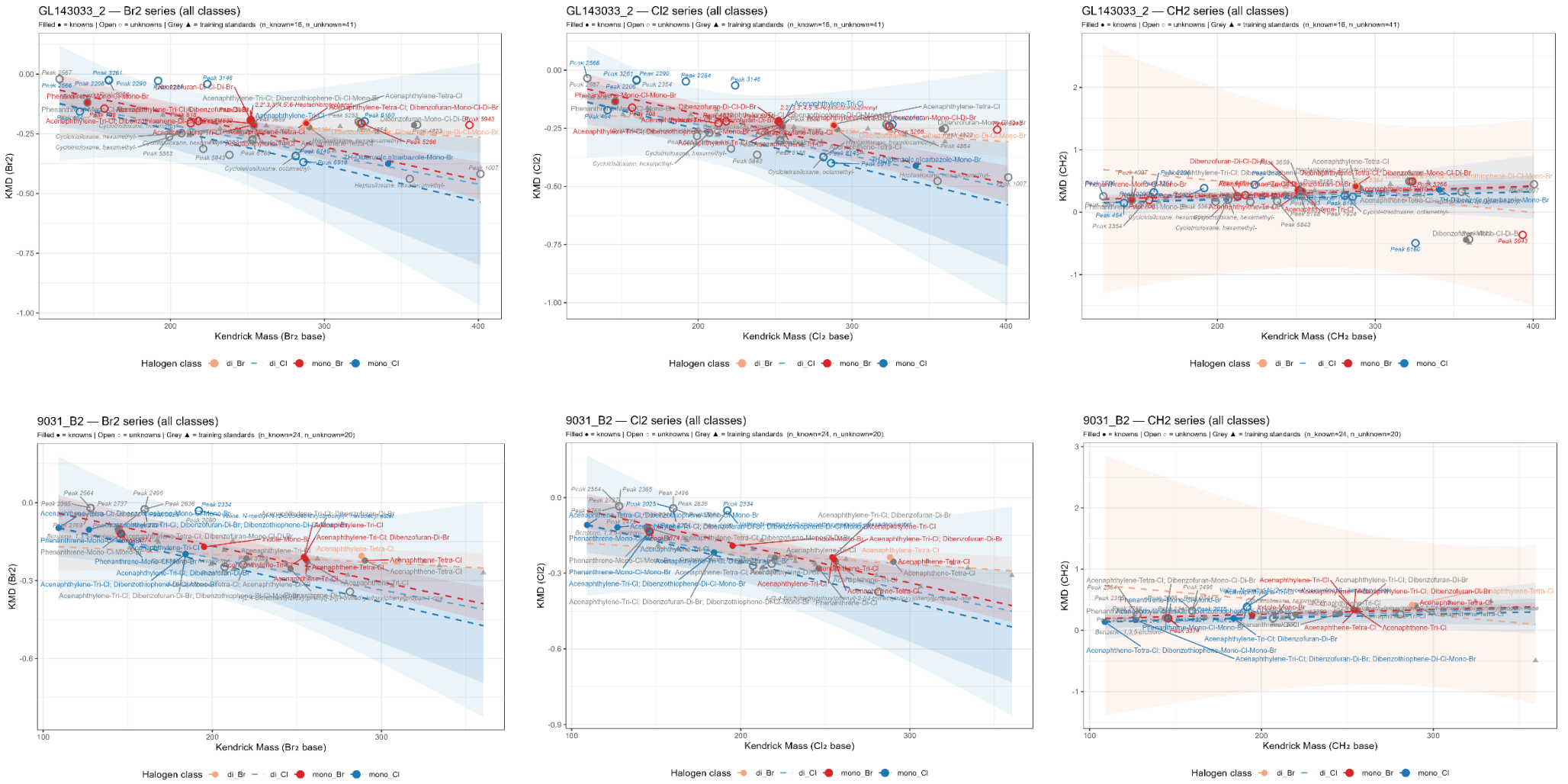



**Figure V 2:** Kendrick mass defect plots generated for the two sediment samples

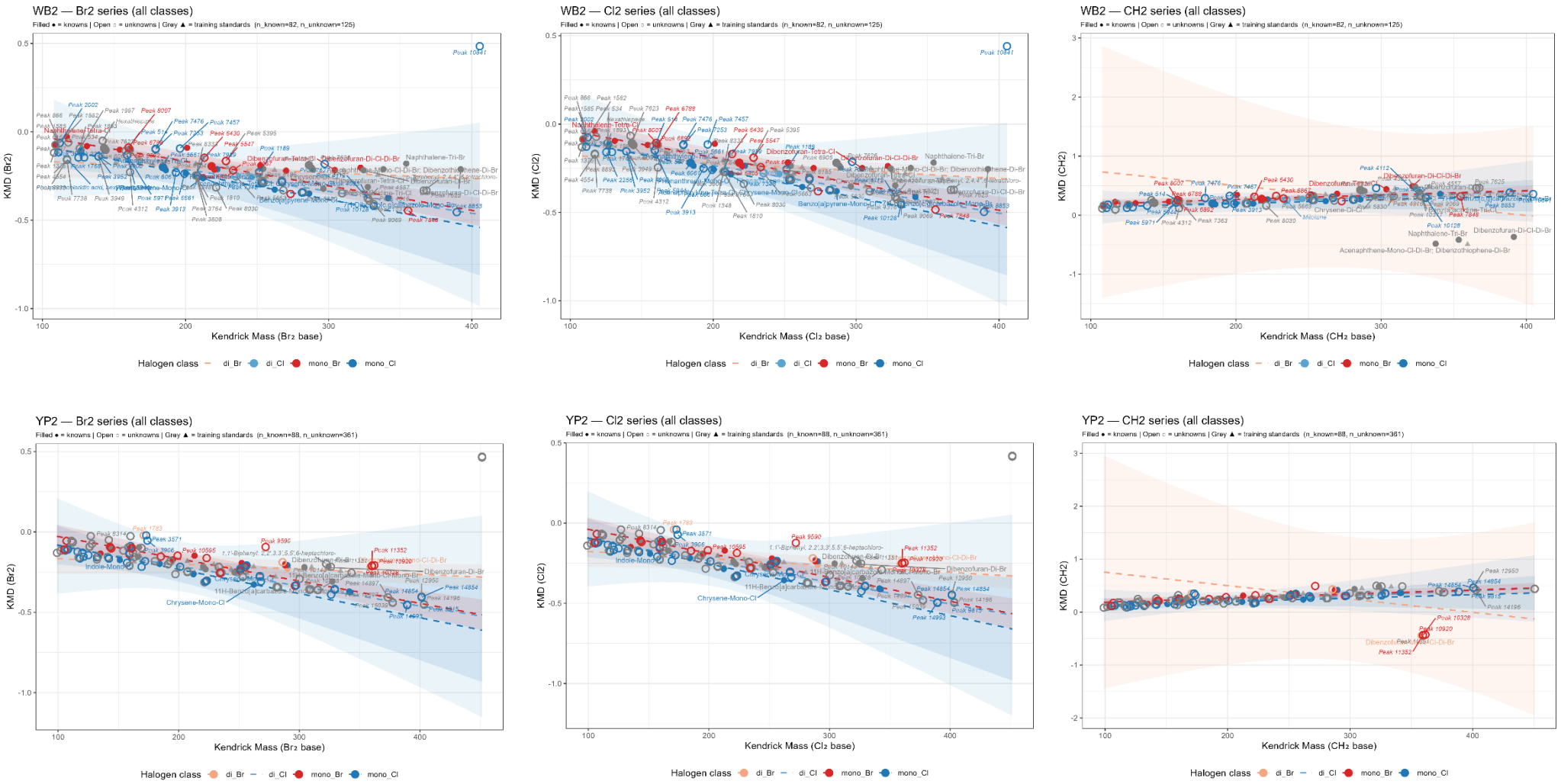



**Figure V 3:** Kendrick mass defect plots generated for the two fish samples

**Table V 9:** Processed output from the KMD assessment of the standard data obtained at workflow 2

| Name from ChromaToF | Halogen class from WF2 step 4 | Possible halo-PAC assignment from WF2 step 5 | km_CH2 | kmd_CH2 | km_Cl2 | kmd_Cl2 | km_Br2 | kmd_Br2 | KMD confidence assignment |
|---|---|---|---|---|---|---|---|---|---|
| 5-Chlorobenzothiophene | mono_Cl | Benzothiophene-Mono-Cl | 167.7921 | 0.207878 | 168.1724 | -0.17245 | 168.1542 | -0.15419 | High at all 3 series within 95% PI |
| Peak 304 | mono_Cl | Quinoline-Mono-Cl | 162.8354 | 0.164597 | 163.2045 | -0.20449 | 163.1868 | -0.18677 | High at all 3 series within 95% PI |
| 2-bromobenzothiophene | mono_Br | Benzothiophene-Mono-Br | 211.6929 | 0.307142 | 212.1727 | -0.17269 | 212.1497 | -0.14965 | High at all 3 series within 95% PI |
| 1,2-Bis-(4-bromophenyl)-N,N,N',N'-tetramethyl-ethane-1,2-diamine | mono_Br | Benzothiophene-Mono-Br | 211.6941 | 0.305904 | 212.1739 | -0.17393 | 212.1509 | -0.15089 | High at all 3 series within 95% PI |
| Peak 383 | mono_Cl; di_Cl | Benzothiophene-Di-Cl | 122.8613 | 0.138691 | 123.1398 | -0.13979 | 123.1264 | -0.12642 | High at all 3 series within 95% PI |
| 3-bromoisoquinoline | mono_Br | Quinoline-Mono-Br | 206.7364 | 0.263582 | 207.205 | -0.20501 | 207.1825 | -0.18252 | Not supported by KMD |
| 2-(p-Chlorophenyl)-3-methylbutyronitrile | mono_Cl | Indole-Mono-Cl | 150.8486 | 0.151437 | 151.1905 | -0.19048 | 151.1741 | -0.17407 | High at all 3 series within 95% PI |
| 3-bromoisoquinoline | mono_Br | Quinoline-Mono-Br | 206.7367 | 0.263322 | 207.2053 | -0.20527 | 207.1828 | -0.18278 | Not supported by KMD |
| 1H-Indole, 5-bromo- | mono_Br | Indole-Mono-Br | 194.7504 | 0.249623 | 195.1918 | -0.1918 | 195.1706 | -0.17061 | High at all 3 series within 95% PI |
| Peak 534 | mono_Br | Indole-Mono-Br | 194.75 | 0.250013 | 195.1914 | -0.19141 | 195.1702 | -0.17022 | High at all 3 series within 95% PI |
| Peak 590 | mono_Br; di_Br; poly_or_mixed | Benzothiophene-Di-Br | 210.6853 | 0.314746 | 211.1628 | -0.1628 | 211.1399 | -0.13987 | High at all 3 series within 95% PI |
| Peak 727 | mono_Br | Acenaphthene-Mono-Br; Dibenzofuran-Di-Cl | 231.727 | 0.272958 | 232.2523 | -0.25228 | 232.2271 | -0.22707 | High at all 3 series within 95% PI |
| 9-bromofluorene | mono_Br | Fluorene-Mono-Br | 243.7158 | 0.284249 | 244.2682 | -0.26816 | 244.2416 | -0.24164 | High at all 3 series within 95% PI |
| 9-Hydroxymethyl-9,10-dihydro-2-chloro acridine | mono_Cl | Acridine-Mono-Cl | 212.7945 | 0.205534 | 213.2768 | -0.27679 | 213.2536 | -0.25364 | High at all 3 series within 95% PI |
| Anthracene, 2-chloro- | mono_Cl | Phenanthrene-Mono-Cl | 211.8016 | 0.198354 | 212.2817 | -0.28172 | 212.2587 | -0.25868 | High at all 3 series within 95% PI |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Peak 971 | mono_Cl; poly_or_mixed | Fluorene-Di-Cl | 198.8077 | 0.192259 | 199.2584 | -0.25837 | 199.2367 | -0.23673 | High at all 3 series within 95% PI |
| Peak 975 | di_Br; poly_or_mixed | Acenaphthylene-Di-Br | 307.5379 | 0.462083 | 308.235 | -0.23499 | 308.2015 | -0.20153 | High at all 3 series within 95% PI |
| dibenzo[b,d]thiophene, 2-bromo- | mono_Br | Dibenzothiophene-Mono-Br | 261.6521 | 0.34791 | 262.2452 | -0.24516 | 262.2167 | -0.21669 | High at all 3 series within 95% PI |
| Fluoranthene | mono_Br | Dibenzothiophene-Mono-Br | 261.6506 | 0.349408 | 262.2437 | -0.24366 | 262.2152 | -0.21519 | High at all 3 series within 95% PI |
| Peak 1043 | mono_Br | Acridine-Mono-Br | 256.6988 | 0.301203 | 257.2806 | -0.28064 | 257.2527 | -0.25271 | High at all 3 series within 95% PI |
| Peak 1048 | mono_Br | Acridine-Mono-Br | 256.6951 | 0.304898 | 257.2769 | -0.27693 | 257.249 | -0.249 | High at all 3 series within 95% PI |
| Peak 1060 | mono_Cl | Carbazole-Mono-Cl | 200.809 | 0.191026 | 201.2641 | -0.26413 | 201.2423 | -0.24228 | High at all 3 series within 95% PI |
| Peak 1064 | mono_Br | Phenanthrene-Mono-Br | 255.7015 | 0.298528 | 256.2811 | -0.28105 | 256.2532 | -0.25323 | High at all 3 series within 95% PI |
| 9-Bromoanthracene | mono_Br | Phenanthrene-Mono-Br | 255.7017 | 0.298268 | 256.2813 | -0.28131 | 256.2535 | -0.25349 | High at all 3 series within 95% PI |
| 1,4-dichloroanthracene | di_Cl; mono_Br; poly_or_mixed | Phenanthrene-Di-Cl | 245.725 | 0.274965 | 246.282 | -0.282 | 246.2553 | -0.25526 | High at all 3 series within 95% PI |
| Peak 1216 | mono_Cl | Dibenzofuran-Di-Cl-Mono-Br | 311.5466 | 0.453375 | 312.2528 | -0.25279 | 312.2189 | -0.21889 | High at all 3 series within 95% PI |
| Peak 1246 | mono_Br; di_Br; poly_or_mixed | Fluorene-Di-Br | 242.7086 | 0.291354 | 243.2588 | -0.25878 | 243.2324 | -0.23237 | High at all 3 series within 95% PI |
| Peak 1489 | mono_Br | Fluoranthene-Mono-Br | 279.6747 | 0.325306 | 280.3086 | -0.30861 | 280.2782 | -0.27818 | High at all 3 series within 95% PI |
| Peak 1584 | di_Br; mono_Cl | Phenanthrene-Di-Br | 333.524 | 0.476002 | 334.28 | -0.27997 | 334.2437 | -0.24368 | High at all 3 series within 95% PI |
| Peak 1595 | mono_Br | Fluoranthene-Mono-Br | 279.6732 | 0.326755 | 280.3072 | -0.30716 | 280.2767 | -0.27673 | High at all 3 series within 95% PI |
| Peak 1882 | mono_Br; di_Br; poly_or_mixed | Carbazole-Di-Br; Dibenzofuran-Mono-Br | 243.7032 | 0.296815 | 244.2556 | -0.25557 | 244.2291 | -0.22905 | High at all 3 series within 95% PI |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 9-Bromo-1,5-chloroanthracene | poly_or_mixed | Phenanthrene-Di-Cl-Mono-Br | 323.5454 | 0.454618 | 324.2787 | -0.27874 | 324.2435 | -0.24353 | High at all 3 series within 95% PI |
| Peak 2394 | mono_Br | Benz(a)acridine-Mono-Br; Chrysene-Mono-Br | 305.6609 | 0.339055 | 306.3538 | -0.35377 | 306.3205 | -0.32051 | High at all 3 series within 95% PI |

**Table V 10:** HPACs reference masses used in workflow 3

| Compound name | Formula | Exact mass |
|---|---|---|
| Indole | C8H7N | 117.0579 |
| Indole-Mono-Cl | C8H6ClN | 151.0189 |
| Indole-Di-Cl | C8H5Cl2N | 184.9799 |
| Indole-Tri-Cl | C8H4Cl3N | 218.9409 |
| Indole-Mono-Br | C8H6BrN | 194.9684 |
| Indole-Di-Br | C8H5Br2N | 272.8789 |
| Indole-Tri-Br | C8H4Br3N | 350.7894 |
| Methyl-Indole | C9H9N | 131.0735 |
| Ethyl-Indole | C10H11N | 145.0892 |
| Propyl-Indole | C11H13N | 159.1048 |
| Butyl-Indole | C12H15N | 173.1205 |
| Quinoline | C9H7N | 129.0579 |
| Quinoline-Mono-Cl | C9H6ClN | 163.0189 |
| Quinoline-Di-Cl | C9H5Cl2N | 196.9799 |
| Quinoline-Tri-Cl | C9H4Cl3N | 230.9409 |
| Quinoline-Mono-Br | C9H6BrN | 206.9684 |
| Quinoline-Di-Br | C9H5Br2N | 284.8789 |
| Quinoline-Tri-Br | C9H4Br3N | 362.7894 |
| Methyl-Quinoline | C10H9N | 143.0735 |
| Ethyl-Quinoline | C11H11N | 157.0892 |
| Propyl-Quinoline | C12H13N | 171.1048 |
| Butyl-Quinoline | C13H15N | 185.1205 |
| Acridine | C13H9N | 179.0735 |
| Acridine-Mono-Cl | C13H8ClN | 213.0345 |
| Acridine-Di-Cl | C13H7Cl2N | 246.9956 |
| Acridine-Tri-Cl | C13H6Cl3N | 280.9566 |
| Acridine-Mono-Br | C13H8BrN | 256.984 |
| Acridine-Di-Br | C13H7Br2N | 334.8945 |
| Acridine-Tri-Br | C13H6Br3N | 412.805 |
| Methyl-Acridine | C14H11N | 193.0892 |
| Ethyl-Acridine | C15H13N | 207.1048 |
| Propyl-Acridine | C16H15N | 221.1205 |
| Butyl-Acridine | C17H17N | 235.1361 |
| Carbazole | C12H9N | 167.0735 |
| Carbazole-Mono-Cl | C12H8ClN | 201.0345 |
| Carbazole-Di-Cl | C12H7Cl2N | 234.9956 |
| Carbazole-Tri-Cl | C12H6Cl3N | 268.9566 |
| Carbazole-Mono-Br | C12H8BrN | 244.984 |
| Carbazole-Di-Br | C12H7Br2N | 322.8945 |

| | | |
|---|---|---|
| Carbazole-Tri-Br | C12H6Br3N | 400.805 |
| Methyl-Carbazole | C13H11N | 181.0892 |
| Ethyl-Carbazole | C14H13N | 195.1048 |
| Propyl-Carbazole | C15H15N | 209.1205 |
| Butyl-Carbazole | C16H17N | 223.1361 |
| Benz(a)acridine | C17H11N | 229.0892 |
| Benz(a)acridine-Mono-Cl | C17H10ClN | 263.0502 |
| Benz(a)acridine-Di-Cl | C17H9Cl2N | 297.0112 |
| Benz(a)acridine-Tri-Cl | C17H8Cl3N | 330.9722 |
| Benz(a)acridine-Mono-Br | C17H10BrN | 306.9997 |
| Benz(a)acridine-Di-Br | C17H9Br2N | 384.9102 |
| Benz(a)acridine-Tri-Br | C17H8Br3N | 462.8207 |
| Methyl-Benz(a)acridine | C18H13N | 243.1048 |
| Ethyl-Benz(a)acridine | C19H15N | 257.1205 |
| Propyl-Benz(a)acridine | C20H17N | 271.1361 |
| Butyl-Benz(a)acridine | C21H19N | 285.1518 |
| 11H-Benzo[a]carbazole | C16H11N | 217.0892 |
| 11H-Benzo[a]carbazole-Mono-Cl | C16H10ClN | 251.0502 |
| 11H-Benzo[a]carbazole-Di-Cl | C16H9Cl2N | 285.0112 |
| 11H-Benzo[a]carbazole-Tri-Cl | C16H8Cl3N | 318.9722 |
| 11H-Benzo[a]carbazole-Mono-Br | C16H10BrN | 294.9997 |
| 11H-Benzo[a]carbazole-Di-Br | C16H9Br2N | 372.9101 |
| 11H-Benzo[a]carbazole-Tri-Br | C16H8Br3N | 450.8207 |
| Methyl-11H-Benzo[a]carbazole | C17H13N | 231.1048 |
| Ethyl-11H-Benzo[a]carbazole | C18H15N | 245.1205 |
| Propyl-11H-Benzo[a]carbazole | C19H17N | 259.1361 |
| Butyl-11H-Benzo[a]carbazole | C20H19N | 273.1518 |
| Dibenz[a,j]acridine | C21H13N | 279.1048 |
| Dibenz[a,j]acridine-Mono-Cl | C21H12ClN | 313.0658 |
| Dibenz[a,j]acridine-Di-Cl | C21H11Cl2N | 347.0268 |
| Dibenz[a,j]acridine-Tri-Cl | C21H10Cl3N | 380.9878 |
| Dibenz[a,j]acridine-Mono-Br | C21H12BrN | 357.0153 |
| Dibenz[a,j]acridine-Di-Br | C21H11Br2N | 434.9258 |
| Dibenz[a,j]acridine-Tri-Br | C21H10Br3N | 512.8363 |
| Methyl-Dibenz[a,j]acridine | C22H15N | 293.1205 |
| Ethyl-Dibenz[a,j]acridine | C23H17N | 307.1361 |
| Propyl-Dibenz[a,j]acridine | C24H19N | 321.1518 |
| Butyl-Dibenz[a,j]acridine | C25H21N | 335.1674 |
| 7H-Dibenzo[c,g]carbazole | C20H13N | 267.1048 |
| 7H-Dibenzo[c,g]carbazole-Mono-Cl | C20H12ClN | 301.0658 |
| 7H-Dibenzo[c,g]carbazole-Di-Cl | C20H11Cl2N | 335.0268 |
| 7H-Dibenzo[c,g]carbazole-Tri-Cl | C20H10Cl3N | 368.9879 |

| | | |
|---|---|---|
| 7H-Dibenzo[c,g]carbazole-Mono-Br | C20H12BrN | 345.0153 |
| 7H-Dibenzo[c,g]carbazole-Di-Br | C20H11Br2N | 422.9258 |
| 7H-Dibenzo[c,g]carbazole-Tri-Br | C20H10Br3N | 500.8363 |
| Methyl-7H-Dibenzo[c,g]carbazole | C21H15N | 281.1205 |
| Ethyl-7H-Dibenzo[c,g]carbazole | C22H17N | 295.1361 |
| Propyl-7H-Dibenzo[c,g]carbazole | C23H19N | 309.1518 |
| Butyl-7H-Dibenzo[c,g]carbazole | C24H21N | 323.1674 |
| Benzothiophene | C8H6S | 134.019 |
| Benzothiophene-Mono-Cl | C8H5ClS | 167.9801 |
| Benzothiophene-Di-Cl | C8H4Cl2S | 201.9411 |
| Benzothiophene-Tri-Cl | C8H3Cl3S | 235.9021 |
| Benzothiophene-Mono-Br | C8H5BrS | 211.9295 |
| Benzothiophene-Di-Br | C8H4Br2S | 289.8401 |
| Benzothiophene-Tri-Br | C8H3Br3S | 367.7506 |
| Methyl-Benzothiophene | C9H8S | 148.0347 |
| Ethyl-Benzothiophene | C10H10S | 162.0503 |
| Propyl-Benzothiophene | C11H12S | 176.066 |
| Butyl-Benzothiophene | C12H14S | 190.0816 |
| Dibenzothiophene | C12H8S | 184.0347 |
| Dibenzothiophene-Mono-Cl | C12H7ClS | 217.9957 |
| Dibenzothiophene-Di-Cl | C12H6Cl2S | 251.9567 |
| Dibenzothiophene-Tri-Cl | C12H5Cl3S | 285.9178 |
| Dibenzothiophene-Mono-Br | C12H7BrS | 261.9452 |
| Dibenzothiophene-Di-Br | C12H6Br2S | 339.8557 |
| Dibenzothiophene-Tri-Br | C12H5Br3S | 417.7662 |
| Methyl-Dibenzothiophene | C13H10S | 198.0503 |
| Ethyl-Dibenzothiophene | C14H12S | 212.066 |
| Propyl-Dibenzothiophene | C15H14S | 226.0816 |
| Butyl-Dibenzothiophene | C16H16S | 240.0973 |
| Benzo(b)naphtho(1,2-d)thiophene | C16H10S | 234.0503 |
| Benzo(b)naphtho(1,2-d)thiophene-Mono-Cl | C16H9ClS | 268.0114 |
| Benzo(b)naphtho(1,2-d)thiophene-Di-Cl | C16H8Cl2S | 301.9724 |
| Benzo(b)naphtho(1,2-d)thiophene-Tri-Cl | C16H7Cl3S | 335.9334 |
| Benzo(b)naphtho(1,2-d)thiophene-Mono-Br | C16H9BrS | 311.9608 |
| Benzo(b)naphtho(1,2-d)thiophene-Di-Br | C16H8Br2S | 389.8714 |
| Benzo(b)naphtho(1,2-d)thiophene-Tri-Br | C16H7Br3S | 467.7819 |
| Methyl-Benzo(b)naphtho(1,2-d)thiophene | C17H12S | 248.066 |
| Ethyl-Benzo(b)naphtho(1,2-d)thiophene | C18H14S | 262.0816 |
| Propyl-Benzo(b)naphtho(1,2-d)thiophene | C19H16S | 276.0973 |
| Butyl-Benzo(b)naphtho(1,2-d)thiophene | C20H18S | 290.1129 |
| Benzofuran | C8H6O | 118.0419 |
| Benzofuran-Mono-Cl | C8H5ClO | 152.0029 |

| | | |
|---|---|---|
| Benzofuran-Di-Cl | C8H4Cl2O | 185.9639 |
| Benzofuran-Tri-Cl | C8H3Cl3O | 219.925 |
| Benzofuran-Mono-Br | C8H5BrO | 195.9524 |
| Benzofuran-Di-Br | C8H4Br2O | 273.8629 |
| Benzofuran-Tri-Br | C8H3Br3O | 351.7734 |
| Methyl-Benzofuran | C9H8O | 132.0575 |
| Ethyl-Benzofuran | C10H10O | 146.0732 |
| Propyl-Benzofuran | C11H12O | 160.0888 |
| Butyl-Benzofuran | C12H14O | 174.1045 |
| Dibenzofuran | C12H8O | 168.0575 |
| Dibenzofuran-Mono-Cl | C12H7ClO | 202.0185 |
| Dibenzofuran-Di-Cl | C12H6Cl2O | 235.9796 |
| Dibenzofuran-Tri-Cl | C12H5Cl3O | 269.9406 |
| Dibenzofuran-Mono-Br | C12H7BrO | 245.968 |
| Dibenzofuran-Di-Br | C12H6Br2O | 323.8785 |
| Dibenzofuran-Tri-Br | C12H5Br3O | 401.7891 |
| Methyl-Dibenzofuran | C13H10O | 182.0732 |
| Ethyl-Dibenzofuran | C14H12O | 196.0888 |
| Propyl-Dibenzofuran | C15H14O | 210.1045 |
| Butyl-Dibenzofuran | C16H16O | 224.1201 |
| Benzonaphthodibenzofuran | C16H10O | 218.0732 |
| Benzonaphthodibenzofuran-Mono-Cl | C16H9ClO | 252.0342 |
| Benzonaphthodibenzofuran-Di-Cl | C16H8Cl2O | 285.9952 |
| Benzonaphthodibenzofuran-Tri-Cl | C16H7Cl3O | 319.9562 |
| Benzonaphthodibenzofuran-Mono-Br | C16H9BrO | 295.9837 |
| Benzonaphthodibenzofuran-Di-Br | C16H8Br2O | 373.8942 |
| Benzonaphthodibenzofuran-Tri-Br | C16H7Br3O | 451.8047 |
| Methyl-Benzonaphthodibenzofuran | C17H12O | 232.0888 |
| Ethyl-Benzonaphthodibenzofuran | C18H14O | 246.1045 |
| Propyl-Benzonaphthodibenzofuran | C19H16O | 260.1201 |
| Butyl-Benzonaphthodibenzofuran | C20H18O | 274.1358 |

**Table V 11**: The performance characteristics of WF3 at different evidence score thresholds

| **Evidence score threshold** | **Features flagged** | **Sensitivity** | **Specificity** | **Precision** | **F1** |
|---|---|---|---|---|---|
| ≥ 3 (Medium and above) | 354 | 1.000 | 0.088 | 0.065 | 0.122 |
| ≥ 5 (High tier, as reported) | 293 | 0.913 | 0.251 | 0.072 | 0.133 |
| ≥ 6 | 128 | 0.478 | 0.678 | 0.086 | 0.146 |
| ≥ 7 | 12 | 0.043 | 0.970 | 0.083 | 0.057 |

**Table V 12**: A comparison of the features and performance characteristics between W2 and WF3

| **Characteristic** | **WF2 (halogen isotope filter)** | **WF3 (NSO heteroatom screen)** |
|---|---|---|
| Diagnostic signature | $^{37}Cl$ / $^{81}Br$ isotopic signatures | Mass defect, neutral losses, DBE |
| Nature of evidence | Physical constant of the nuclide | Inferential and shared with many organics |
| Sensitivity | 0.955 | 0.913 (at high tiers) |
| Specificity | 0.841 | 0.251 |
| F1 | 0.898 | 0.133 |
| Enrichment over base rate | 1.71× | 1.20× |
| Intended role | Near decisive class assignment | High recall candidate generation |

**Table V 13:** Compounds detected in the environmental samples

| Species | Targeted PACs | | High confidence non-targeted PACs | | Moderate confidence non-target PACs. | Potential halo-PACs |
|---|---|---|---|---|---|---|
| | **Two or more samples** | **One sample** | **Two or more samples** | **One sample** | **Two or more samples** | **Two or more samples** |
| Yellow Perch | 2,3-Br2-BT<br>Ind<br>2-MeNap<br>1-MeNap<br>1,4-DiMeNap<br>Ace<br>DBF<br>DBT<br>Phe<br>Car<br>2-MeDBT<br>3-MePhe<br>2,6-DiMePhe<br>Flt<br>Pyr<br>9,10-Br2-Ant<br>Chr<br>BaAnt<br>11H-BaCar<br>IcdP | IQui<br>5-Cl-BT<br>Flu<br>2-MeDBF<br>1,4,6,7-TetraMeNap<br>2-MePhe<br>1-MePhe<br>1,3-DiMePhe<br>1,7-DiMePhe<br>9-Br-1,5-Cl-Ant<br>BkF<br>Per<br>BghiP | Nap<br>3-MeIQui<br>1,4-DiMeNap (six isomers)<br>2,3,5-TriMeNap (three isomers)<br>5-Cl-Ind<br>5-Br-Ind (two isomers)<br>9-Br-Flu (two isomers)<br>1,3-DiMePhe<br>2,4,5-TriMeDBT<br>1,2,6-TriMePhe (two isomers)<br>1,4-Cl2-Ant<br>Ret<br>9-Br-1,5-Cl-Ant (two isomers)<br>Per | 1,4,6,7-TetraMeNap<br>1-MeFlu<br>9-Cl-Flu<br>2,4-DiMeDBT<br>1-Cl-Ant<br>BbNF(2,3-d)<br>1-MePyr<br>4,5-DiMePyr (two isomers)<br>1,2,6,9-TetMePhe<br>BbNT(1,2-d)<br>6-MeBbNT(2,1-d)<br>BbF | Qui<br>Phe<br>Ant<br>Car<br>9-Br-Flu<br>Pyr<br>7H-BcCar | Cl2-Nap<br>Cl-BT<br>Br2-DBF<br>Cl3-DBF<br>Cl2-Phe<br>Cl2-BNT<br>Br-7H-DiBenzCar<br>Cl-Chr<br>Br2-DBT<br>Cl2-B(g,h,i)Pyr |
| White Bass | DBF<br>Flu<br>Phe<br>BaAnt | 2-MeNap<br>1-MeNap<br>Acy<br>Ace<br>2-MeDBF<br>DBT<br>Car<br>3-MePhe<br>9-MePhe<br>1-MePhe<br>1,3-DiMePhe<br>Flt<br>Pyr<br>3,6-Cl2-Car<br>9,10-Br2-Ant | Nap<br>1-Cl-Ant<br>1,4-Cl2-Ant (two isomers)<br>Ret | IQui<br>1,4-DiMeNap<br>2,3,5-TriMeNap<br>4-Br-BaA<br>BbF (three isomers) | Ind<br>Phe<br>9-Cl-Flu<br>Car<br>Ant<br>3-MeBfQui<br>9-Br-Flu<br>1,2-Br2-Acy<br>5,6-Br2-1,2-H2-Acy | 1,4-Cl2-Ant (three isomers)<br>1-Cl-Ant<br>3,6-Cl2-Car (two isomers)<br>4-Br-BaA<br>Cl-Chr<br>Br3-DBF<br>Cl4-Nap |
| HH_sed1 | BbT<br>1,4-DiMeNap<br>Flu<br>1-MeFlu<br>Phe<br>2-MePhe<br>1-MePhe<br>2,6-DiMePhe | Flt | 1,4-DiMeNap (two isomers)<br>Acy<br>Flu<br>2-MeDBF<br>1,4,6,7-TetraMeNap<br>1-MeFlu<br>Car<br>2-MePhe | Ace<br>3-MeDBT<br>Flt-d10<br>11H-BaCar<br>BbNT(1,2-d)<br>IcdP<br>BghiP | Nap<br>DBT<br>Ant<br>Acr<br>4,5-DiMePyr<br>Ret<br>4,6-DiEtDBT<br>1,2,6,9-TetMePhe | 1,4-Cl2-Ant |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 1,3-DiMePhe<br>1,7-DiMePhe<br><br>2,4,5-TriMeDBT<br>Pyr<br><br>BaAnt<br><br>BaAnt<br>6-MeBbNT(2,1-d)<br>9-Br-1,5-Cl-Ant<br>BkF<br>Per | | 2,6-DiMePhe<br>1,7-DiMePhe (four isomers)<br>Pyr (two isomers)<br>2,4,5-TriMeDBT (five isomers)<br>BbNF(2,3-d) (three isomers)<br>4,5-DiMePyr<br>1,2,6,9-TetMePhe<br><br>Chr<br>Per | | 7H-BcCar<br>11H-BaCar<br><br>DNF[1,2;1,2-d] | |
| HH_sed2 | BbT<br>Ind<br>1-MeNap<br><br>1,4-DiMeNap<br><br>Acy<br>Ace<br>DBF<br>Flu<br><br>2-MeDBF<br>1,4,6,7-TetraMeNap<br>DBT<br>Ant<br>Car<br>3-MeDBT<br>3-MePhe<br><br>2-MePhe<br><br>2,4-DiMeDBT<br>2,6-DiMePhe<br>1,8-DiMePhe<br><br>Pyr<br><br>BcAcr<br>9,10-Br2-Ant<br>Chr<br>BaAnt<br>BkF<br>Per | 2-MeNap<br>4-MeBT<br>2,3-DiMeBT<br><br>1,4-DiMeNap<br><br>2,3,5-TriMeNap<br>1-MeFlu<br>9-PropFlu<br>Phe<br><br>2-MeDBT<br>9-ButFlu<br><br>9-Br-Flu<br>2,4-DiMeBhQ<br>4-PropDBT<br>Flt<br>2,4,5-TriMeDBT<br>11H-BaCar<br><br>9-Br-1,5-Cl-Ant<br>DNF[1,2;1,2-d]<br>BghiP | Nap<br>Qui<br>1,4-DiMeNap (five isomers)<br>2,3,5-TriMeNap (five isomers)<br>2,3,4-TriMeBT<br>2,3,4-TriMeBT<br>Flu (two isomers)<br>1,4,6,7-TetraMeNap (seven isomers)<br>9-EtFlu<br>1-MeFlu<br><br>2-MeDBF<br>Phe<br>9-PropFlu (four isomers)<br>9-MePhe<br>Pyr<br><br>4,6-DiEtDBT (seven isomers)<br>BbNF(2,3-d)<br>Ret<br>4,5-DiMePyr (two isomers)<br>1,2,6,9-TetMePhe (two isomers) | 5-Cl-Ind<br>2-MeDBF<br>5-Br-Ind<br><br>1-Br-Pyr<br><br>BcAcr<br>BaAnt<br>BbF<br>DNF[1,2;1,2-d]<br><br>7,8-DiMeBaP | IcdP<br>DBahA<br>Acr<br><br>9-PropFlu<br><br>9-Cl-Flu<br>9-ButFlu<br>7H-DiBenzCar<br>7,8-DiMeBaP<br><br>4-Br-BaA | 5-Cl-Ind<br>1,2-Br2-Acy<br>1,4-Cl2-Ant<br><br>Carbazole-Di-Cl |

**Table V 14:** Concentrations of non-targeted analytes calculated using the relative response of neighboring target analytes in sediment samples

| Non targets | RT1 | Area | Response factor (RF) | Concentration in ng/g |
|---|---|---|---|---|
| **HH_Sed1** | | | | |
| 1,4-DiMeNap | 1306 | 521729 | 0.35 | 252 |
| 1,4-DiMeNap | 1363 | 1924619 | 1.28 | 930 |
| Flu | 1723 | 2314681 | 2.81 | 894 |
| 2-MeDBF | 1804 | 455883 | 0.34 | 1202 |
| 1,4,6,7-TetraMeNap | 1861 | 197482 | 0.15 | 521 |
| 1-MeFlu | 1966 | 350608 | 0.19 | 1392 |
| 1-MeFlu | 1984 | 680343 | 0.5 | 1794 |
| 2-MePhe | 2383 | 1106945 | 0.6 | 4396 |
| 2,6-DiMePhe | 2590 | 491051 | 0.26 | 1950 |
| 2,6-DiMePhe | 2593 | 620801 | 0.46 | 1637 |
| 1,7-DiMePhe | 2614 | 2095615 | 1.54 | 5527 |
| 1,7-DiMePhe | 2632 | 1241363 | 0.91 | 3274 |
| 1,8-DiMePhe | 2650 | 449427 | 0.24 | 1785 |
| 1,8-DiMePhe | 2686 | 252131 | 0.08 | 147 |
| 2,4,5-TriMeDBT | 2692 | 154156 | 0.14 | 56 |
| 2,4,5-TriMeDBT | 2704 | 197968 | 0.06 | 115 |
| 2,4,5-TriMeDBT | 2710 | 146682 | 0.05 | 85 |
| 2,4,5-TriMeDBT | 2725 | 204512 | 0.19 | 75 |
| BbNF(2,3-d) | 2728 | 2252162 | 2.09 | 825 |
| 2,4,5-TriMeDBT | 2746 | 297257 | 0.1 | 173 |
| BbNF(2,3-d) | 2749 | 1086858 | 0.35 | 633 |
| 1,2,6-TriMePhe | 2764 | 1925644 | 0.62 | 1121 |
| 2,4,5-TriMeDBT | 2773 | 67835 | 0.02 | 39 |
| 2,4,5-TriMeDBT | 2776 | 150540 | 0.05 | 88 |
| 1,4-Cl2-Ant | 2782 | 12018 | 0.01 | 32 |
| 1,2,6-TriMePhe | 2785 | 1108300 | 0.36 | 645 |
| BbNF(2,3-d) | 2791 | 560878 | 0.18 | 326 |
| 4,5-DiMePyr | 2845 | 3863331 | 1.24 | 2248 |
| Ret | 2866 | 398332 | 0.13 | 232 |
| 1,2,6,9-TetMePhe | 2908 | 879480 | 0.28 | 512 |
| 1,2,6,9-TetMePhe | 2926 | 233740 | 0.08 | 136 |
| **HH_Sed2** | | | | |
| BbT | 901 | 382475 | 0.6 | 9.4 |
| 1,4-DiMeNap | 1309 | 289028 | 0.3 | 3.0 |
| 1,4-DiMeNap | 1357 | 2052755 | 2.0 | 21.3 |
| 1,4-DiMeNap | 1363 | 1270833 | 1.3 | 13.2 |
| 1,4-DiMeNap | 1495 | 56017 | 0.1 | 0.6 |
| 1,4-DiMeNap | 1504 | 280966 | 0.3 | 2.9 |
| 2,3,5-TriMeNap | 1519 | 45832 | 0.0 | 0.5 |
| 2,3,5-TriMeNap | 1549 | 209881 | 0.2 | 2.2 |
| 2,3,4-TriMeBT | 1576 | 10331 | 0.0 | 0.1 |

| | | | | |
|---|---|---|---|---|
| 2,3,4-TriMeBT | 1579 | 11334 | 0.0 | 0.1 |
| 2,3,5-TriMeNap | 1633 | 485642 | 0.5 | 5.0 |
| Flu | 1807 | 42135 | 0.0 | 0.3 |
| Flu | 1810 | 22465 | 0.0 | 1.9 |
| 1,4,6,7-TetraMeNap | 1840 | 9314 | 0.0 | 0.2 |
| 1-MeFlu | 1870 | 26765 | 0.0 | 0.6 |
| 1,4,6,7-TetraMeNap | 1888 | 33203 | 0.1 | 0.7 |
| 1,4,6,7-TetraMeNap | 1900 | 176746 | 0.3 | 3.8 |
| 9-EtFlu | 1933 | 109938 | 0.2 | 2.4 |
| 1-MeFlu | 1957 | 270969 | 0.4 | 5.9 |
| 2-MeDBF | 1966 | 68208 | 0.1 | 1.5 |
| 1,4,6,7-TetraMeNap | 1978 | 53774 | 0.1 | 1.2 |
| 1,4,6,7-TetraMeNap | 1990 | 24222 | 0.0 | 0.5 |
| 1-MeFlu | 1999 | 130131 | 0.2 | 2.8 |
| Phe | 2023 | 12948 | 0.0 | 0.3 |
| 9-PropFlu | 2026 | 55236 | 0.1 | 1.1 |
| 1-MeFlu | 2056 | 11932 | 0.0 | 0.2 |
| 1-MeFlu | 2062 | 18438 | 0.0 | 0.4 |
| 9-ButFlu | 2317 | 9805 | 0.0 | 0.2 |
| 9-ButFlu | 2320 | 13827 | 0.0 | 0.3 |
| 9-ButFlu | 2347 | 72695 | 0.1 | 1.5 |
| 9-Br-Flu | 2350 | 39358 | 0.0 | 0.8 |
| 9-MePhe | 2389 | 255189 | 0.3 | 5.2 |
| Pyr | 2704 | 10820 | 0.0 | 0.4 |
| 4,6-DiEtDBT | 2749 | 28095 | 0.0 | 0.9 |
| 4,6-DiEtDBT | 2770 | 25057 | 0.0 | 0.8 |
| BbNF(2,3-d) | 2782 | 169455 | 0.2 | 5.6 |
| 4,6-DiEtDBT | 2794 | 27917 | 0.0 | 0.9 |
| 4,6-DiEtDBT | 2806 | 25137 | 0.0 | 0.8 |
| 4,6-DiEtDBT | 2809 | 14542 | 0.0 | 0.5 |
| Ret | 2815 | 97108 | 0.1 | 3.2 |
| 4,5-DiMePyr | 2818 | 1750238 | 1.8 | 58.3 |
| 4,6-DiEtDBT | 2827 | 14926 | 0.0 | 0.5 |
| 4,6-DiEtDBT | 2830 | 11698 | 0.0 | 0.4 |
| 4,6-DiEtDBT | 2863 | 6235 | 0.0 | 0.2 |
| 1,2,6,9-TetMePhe | 2872 | 11050 | 0.0 | 0.4 |
| 1,2,6,9-TetMePhe | 2899 | 23480 | 0.0 | 0.8 |
| 4,5-DiMePyr | 2929 | 58203 | 0.1 | 1.9 |

**Evaluation of the applicability of PAH-Finder for the current study**

PAH-Finder (Zhang et al., 2025) was applied without modification to authentic PAC and HPAC standards acquired on both the GC×GC–HRTOF–MS platform used throughout this study and a conventional single-column GC–HRTOF–MS platform. PAH-Finder defines a suspect spectrum as one satisfying both a random forest classification of fragment distribution and the detection of doubly charged fragments. Applied to 127 manually verified standard spectra from the two-dimensional platform, 12 suspects were returned; applied to 177 spectra from the one-dimensional platform with background spectra removed, 12 suspects were returned from 123 PAC-like compounds. The sensitivity was 9.4% and 9.8% respectively that is consistent across platforms.

Diagnostic examination identified the doubly charged fragment criterion as the gating step on both platforms: only 16 of 127 two-dimensional spectra (12.6%) and 22 of 123 one-dimensional PAC-like spectra (17.9%) satisfied it. This criterion requires detection of the $^{13}C$ isotope of a doubly charged fragment, which occurs at approximately 1–4% of base peak intensity in the present spectra, and applies its own internal filter retaining only fragments above 1% of base peak; 33–40% of the doubly charged isotope signals present in the data fall below this value. To rule out export configuration as the cause, the two-dimensional spectra were re-exported at a five-fold lower intensity threshold, which produced no change in the pass rate. The same internal threshold removes 63% of the corresponding signals in the demonstration dataset distributed with PAH-Finder itself, indicating that the constraint is intrinsic to the criterion rather than specific to the present platform.

These observations indicate that PAH-Finder's doubly charged fragment criterion achieves a consistent sensitivity ceiling of approximately 10–20% on spectra with these intensity

characteristics, irrespective of chromatographic platform, while its random forest component behaved as published on the spectra that reached it. This comparison illustrates the platform dependence of identification criteria built on low-abundance diagnostic ions and motivates the evidence streams adopted in the present framework with exact mass, halogen isotope patterns, and homologous-series relationships which remain observable at the intensity levels routinely retained under both one- and two-dimensional acquisition.

Of the twelve suspects PAH-Finder returned from the one-dimensional platform, four formula predictions were additionally identified as unreliable and should not be cited as valid identifications, summarized below.

| **Suspect** | **Predicted formula** | **Correct formula** | **Reason** |
|---|---|---|---|
| **Pyrene, 4,5-dimethyl-** | C17H11 | C18H14 | C17H11 is an odd-electron radical; physically impossible as a stable neutral molecule |
| **Benz[a]anthracene-d12** | C10H20N6O | C18D12 | Rdisop does not handle deuterium; complete misassignment |
| **Benzo[a]pyrene-d12** | C12H20N6O | C20D12 | Same deuterium limitation |
| **Benzo[b]naphtho[1,2-d]thiophene** | C12H13NPS | C16H10S | Spurious N and P; compound contains only C, H, and S |