



PFAS Chemometrics - Forensic Case Studies and Real-World Applications

Zachary Neigh, AECOM



Agenda

- Forensic Concepts
- PFAS Profiling Methods and Applications
- Case Study Highlights

Environmental Forensics: Using Data to Reconstruct a Story

What is Environmental Forensics?

A multidisciplinary practice that investigates the sources, extent, fate, and transport of environmental contaminants. Emerging applications of statistics and machine learning add unrealized value to environmental chemistry datasets.

Example Forensics Use Cases:



Source Identification

- Detected at a site with no history of PFAS usage
- Water Utility detections and residential groundwater wells
- Presence in biological samples



Delineation/Differentiation

- PFAS plume does not follow the expected trend
- Commingled sources must be differentiated and delineated
- Confirmation of offsite fingerprints to source data



Fate and Transport Mechanism Identification

- Chemical evidence of groundwater/surface water interactions
- Influence of historical injections and non-PFAS remediation
- Precursor degradation

Fate and Transport Connections are Complex

Generally, Fate and Transport describes the behavior of chemicals following their release to the environment. This includes the physical, chemical, and biological processes that influence distribution of PFAS in various media, as well as the extent of migration within and between media.



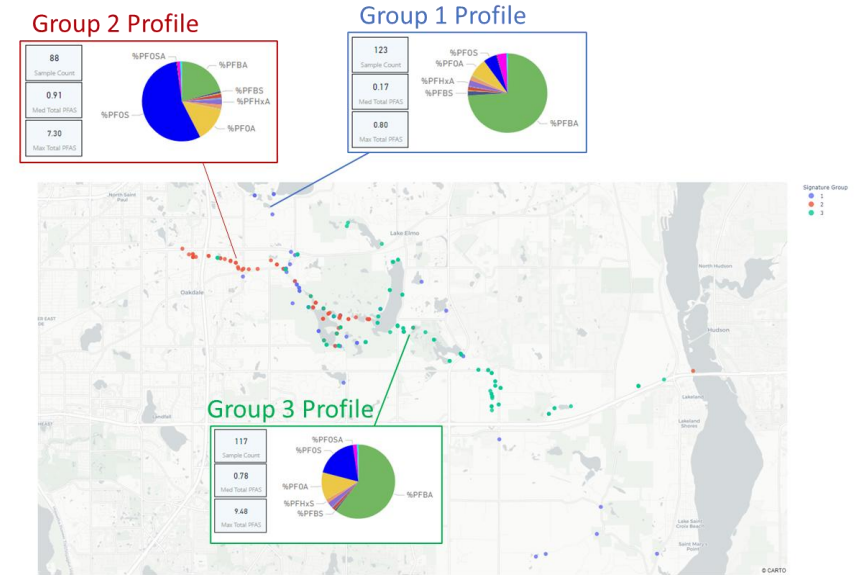
PFAS Fingerprinting versus Profiling

- Human fingerprints are very useful in criminal forensics because they are uniquely tied to one suspect, and they are unchanging over time and distance. Proportions of mixed PFAS in the environment do not act as reliably as fingerprints on their own, because they change during fate and transport processes.
- A forensic profile is a broader classification that incorporates multiple lines of evidence and a range of chemometrics patterns. Including:
 - Spatial metrics – distance from source, gradient, directionality, areas of concern, etc
 - Environmental Media and Interaction Properties – soil type, TOC, grain size, water quality parameters
 - Chemical Properties – branched and linear isomers, chain length, precursor and degradant molar ratios
 - Reference Library Matches – find similar samples or groups of samples in existing databases

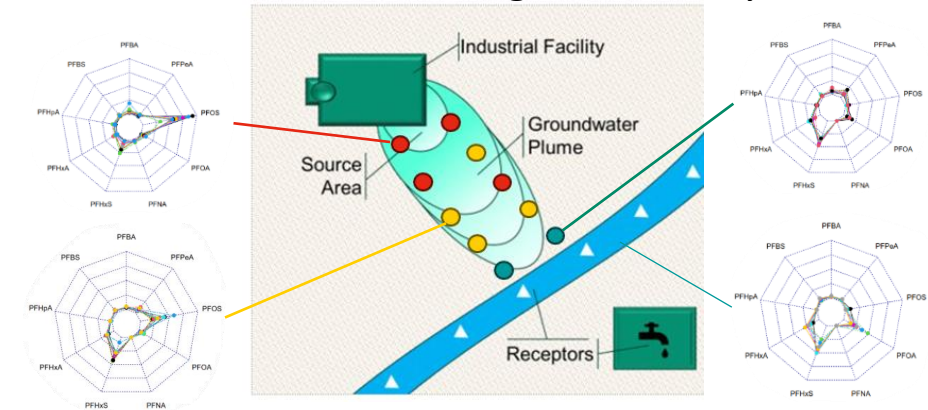
Chemometrics for Environmental Forensics

- Chemometrics is the application of multivariate statistical methods for pattern recognition and machine learning to aid our understanding of chemical data from environmental samples
- Useful as a sophisticated data slicer, detangler, and organizational tool
- Can be used to examine contaminant origins, apportion sources, and understand environmental processes affecting pollutants
- Adapts to different scales easily. We have to specify the right method parameters and levels of cluster sensitivity (i.e. count/variability of samples within each cluster) to tailor it to each case

Large Scale – Low Sensitivity

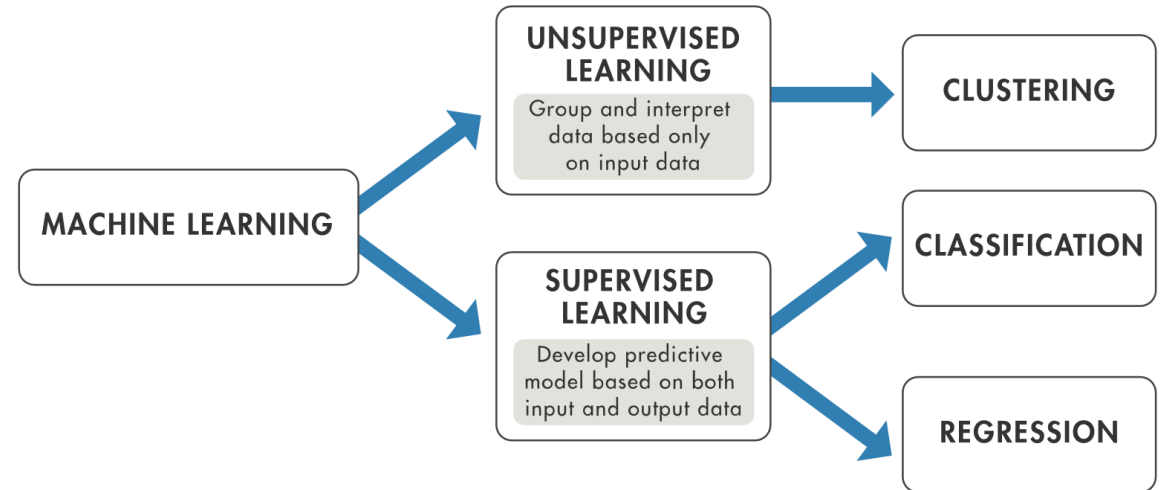
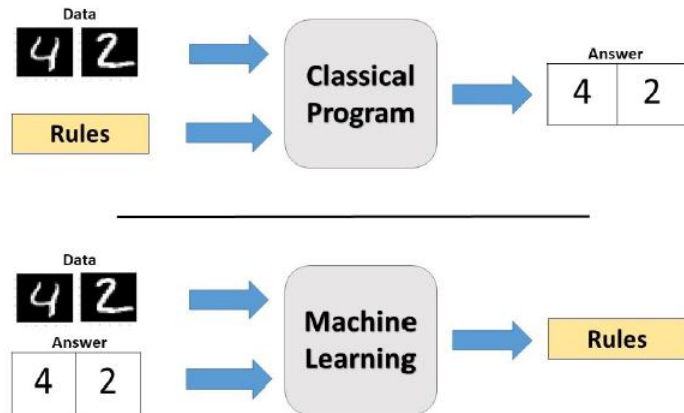


Small Scale – High Sensitivity



Artificial Intelligence (AI) and Machine Learning (ML) in Forensics

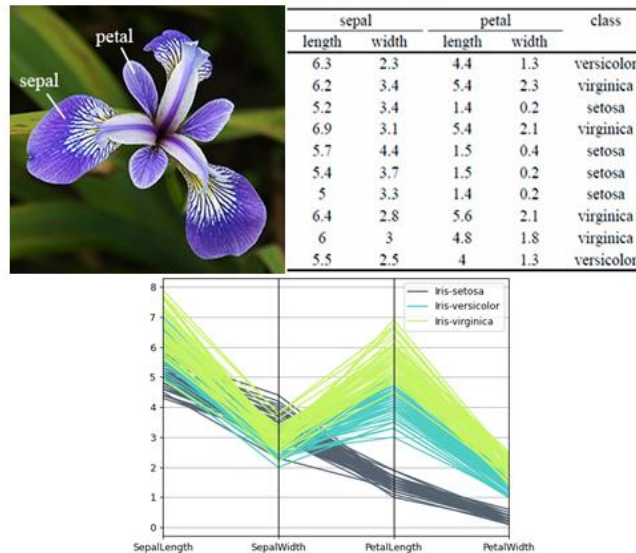
- AI is the ability of a computer to “think” like a human and provide readily interpretable information from data-rich inputs
- ML is a more specific part of AI, where statistical algorithms (i.e. chemometrics) discern the patterns and equations that enable insights and predictions to be made from the data



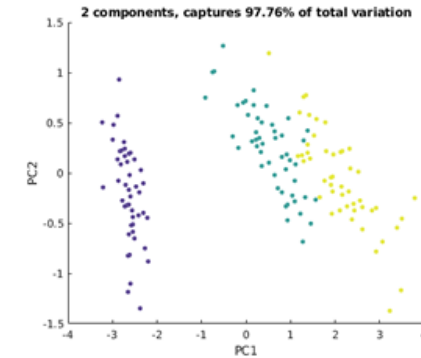
Unsupervised Machine Learning Concept

- A very powerful tool when you have multidimensional data and can't easily see the patterns
- Our human perception is limited to 3 dimensions
- Key Steps to Pattern Recognition
 - Data Preprocessing
 - Dimensionality Reduction
 - Clustering
- More advanced methods can be tailored to specific situations and data types

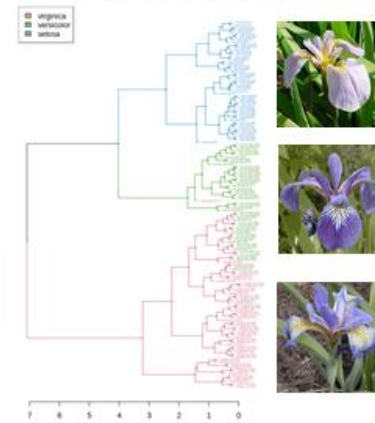
Classic Iris Data Example



Principal Component Analysis



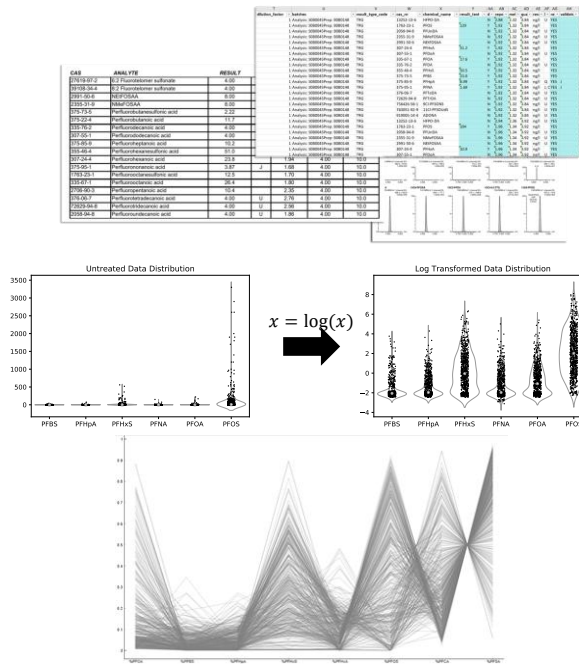
Cluster Analysis and Classification



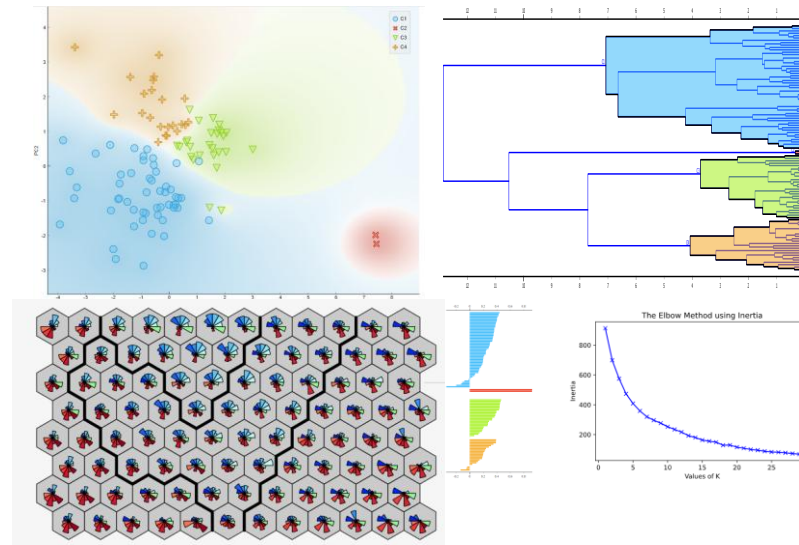
Application to PFAS Data

- Chemical data of all kinds can be broadly explored with this methodology
- Emerging concerns around PFAS and growing complexity of sources have driven development
- Pattern recognition groups the samples by similarity in mixture composition

Data Preprocessing



Pattern Recognition



Expert Interpretation



PFAS Profiling: Data-Driven Forensics Approach

Goals of PFAS Profiling Process

- Tailored PFAS Pattern Screening
 - Forms the foundation of PFAS profiles
 - Provides a roadmap for further investigation
- Profile Building
 - Explain groups of samples, relationships, and significant compositions
 - Identify additional statistical analyses and data needs
- Enhanced Differentiation Techniques
 - Focused analytical and sampling design
 - Specialized Statistical Analysis
 - Improve CSM and drive investigation forward with multiple lines of evidence

Tailored PFAS Pattern Screening

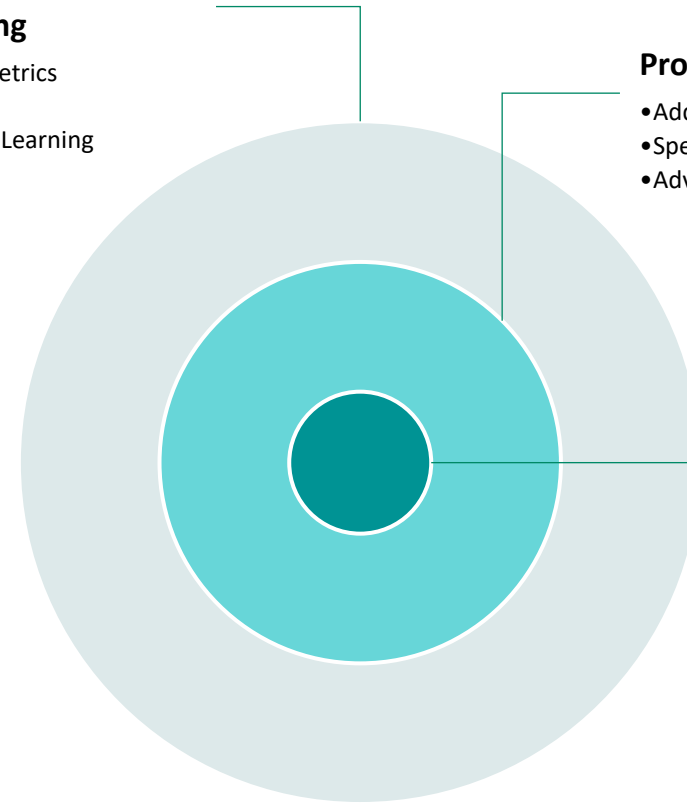
- Chemometrics
- GIS
- Machine Learning

Profile Building

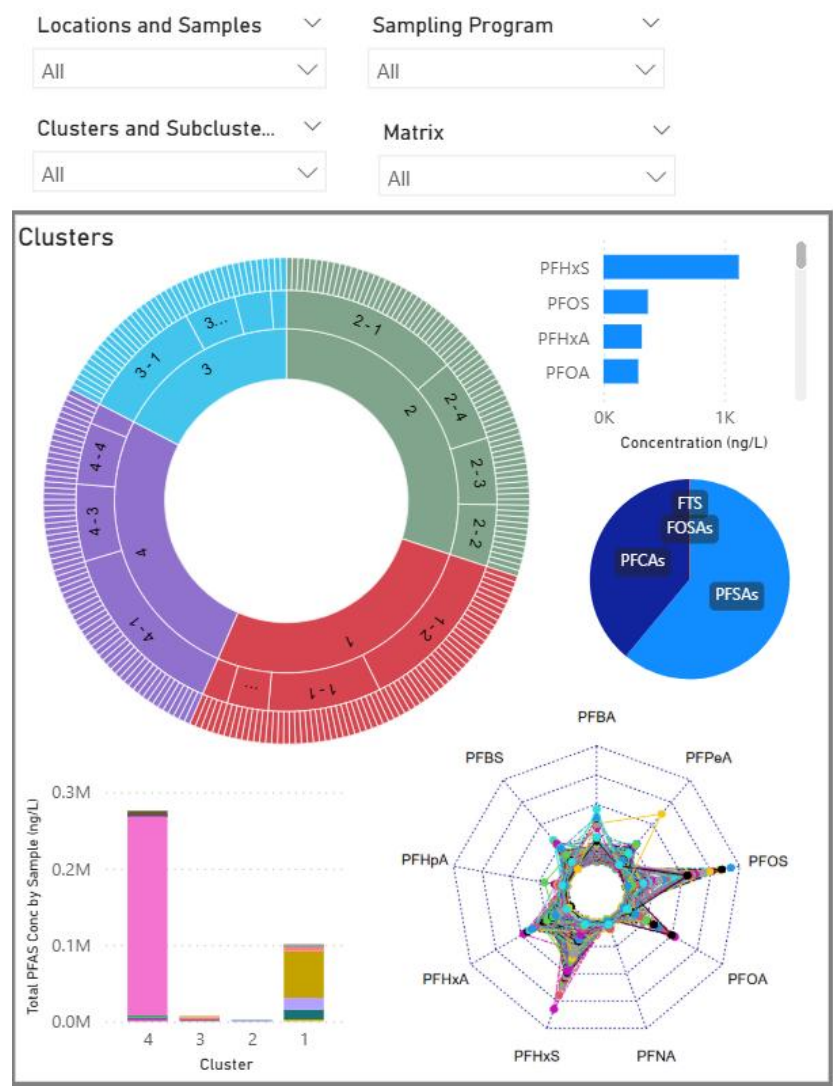
- Additional Data Gathering
- Specialized PFAS Analyses
- Advanced Statistical Analysis

Enhanced Differentiation Techniques

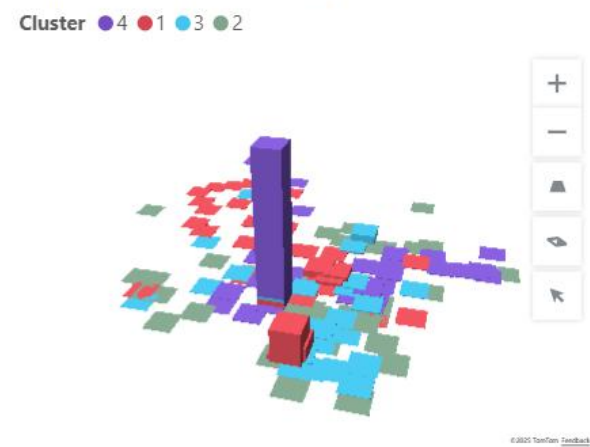
- Non-target analysis
- Isomer quantitation
- Indicator compounds
- Source Apportionment



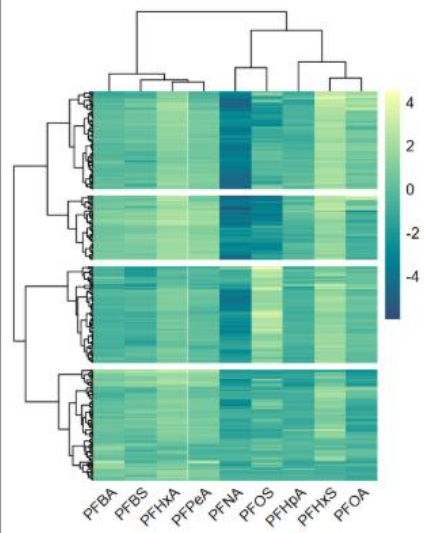
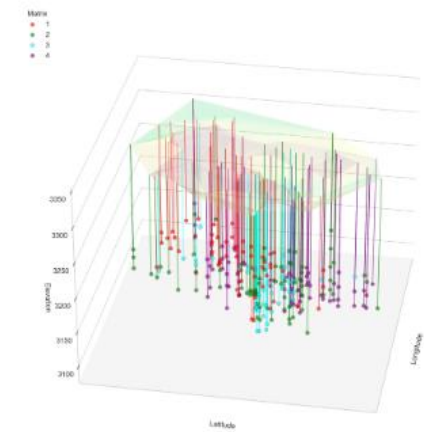
PFAS Chemistry Pattern Screening Example



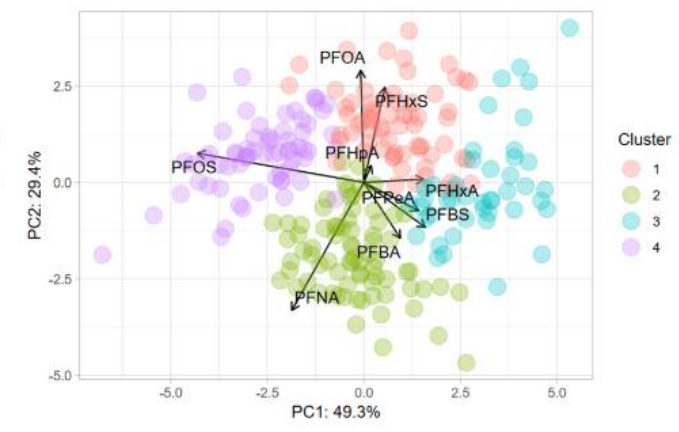
Map of Clusters and Average PFAS Concentration



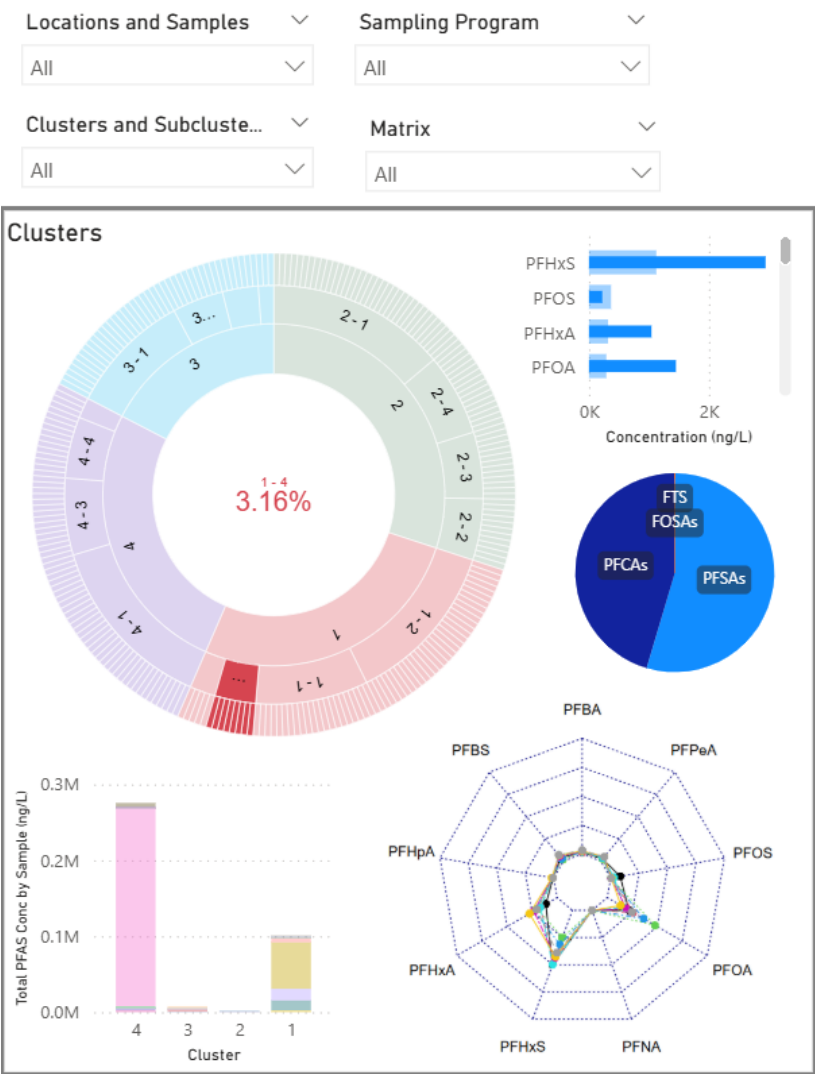
3D Well Cluster Plot



PCA Biplot

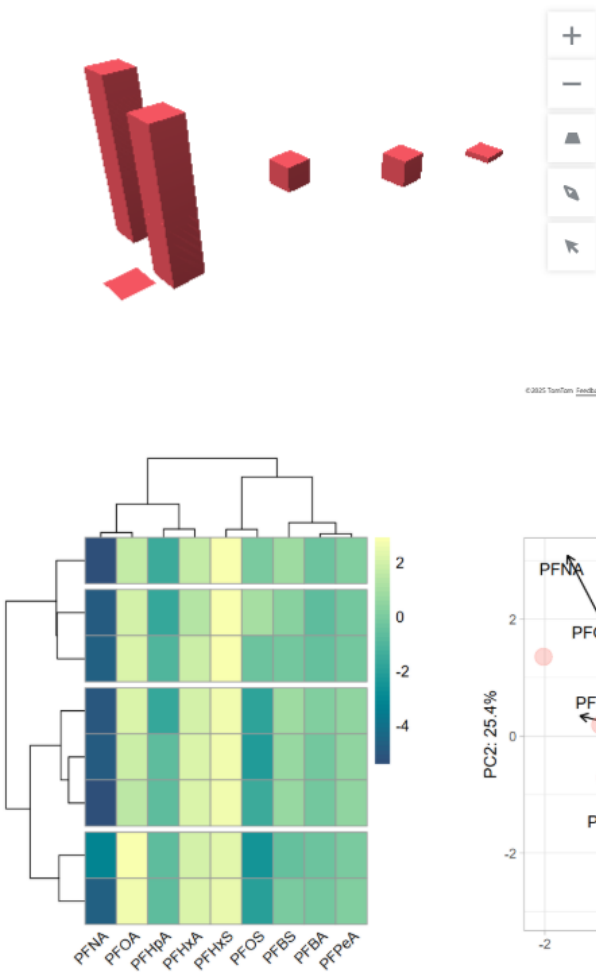


PFAS Chemistry Pattern Screening Example

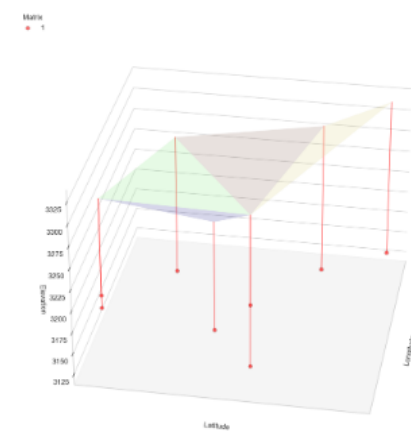


Map of Clusters and Average PFAS Concentration

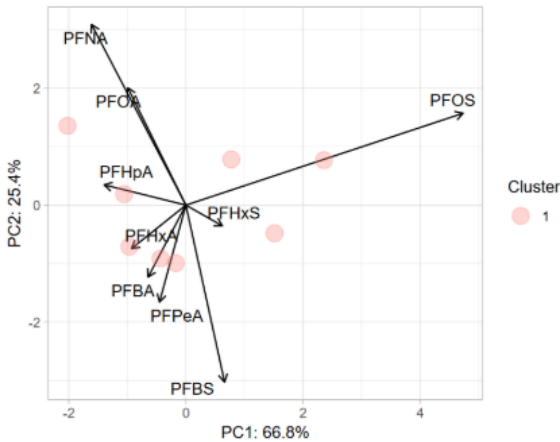
Cluster 1



3D Well Cluster Plot



PCA Biplot



Using Clusters to Identify and Isolate Sources

Cluster 1: Fire Training activities (non-client related)

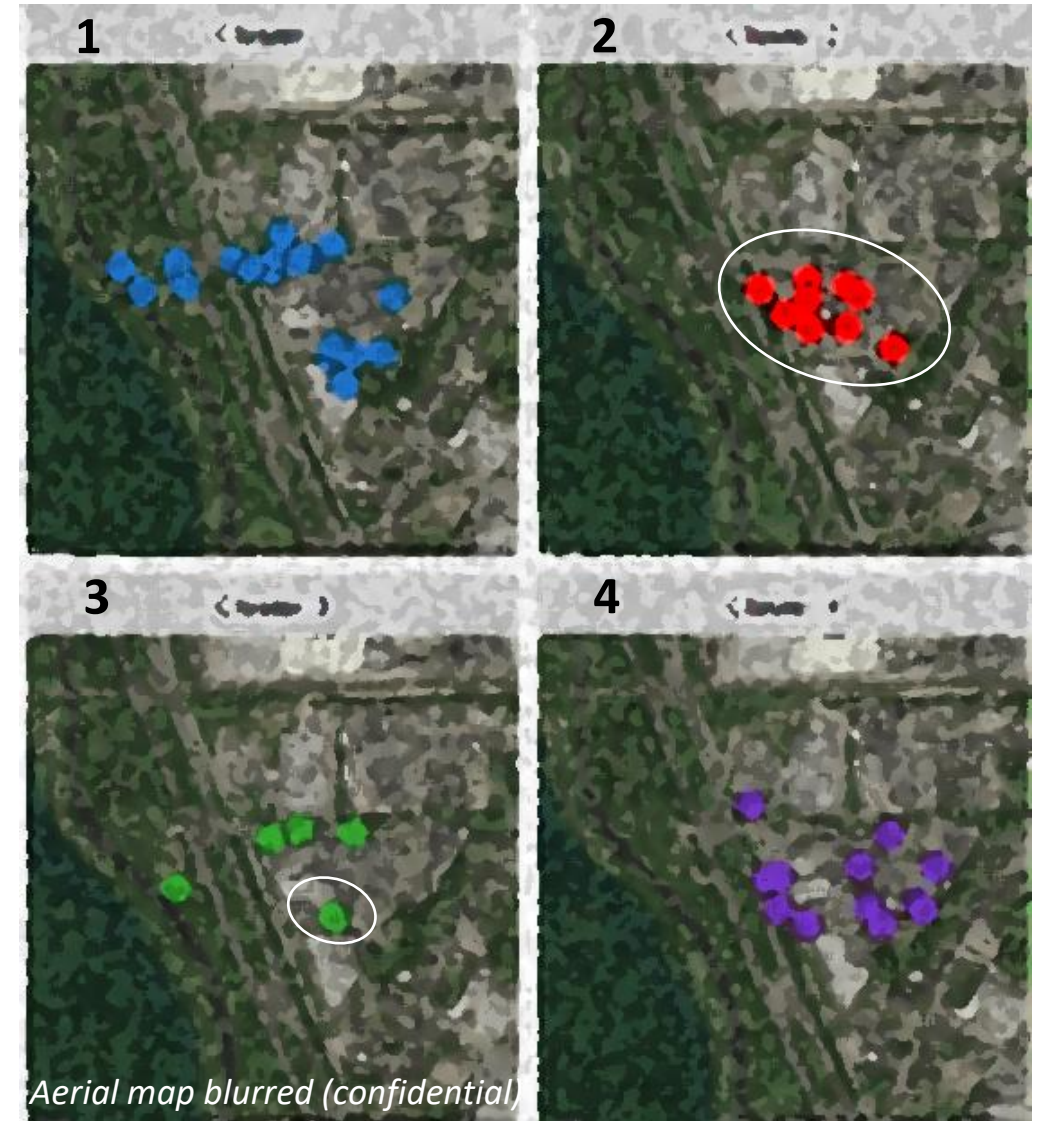
Cluster 2: PFCA - Client related (production)

Cluster 3: PFSA - Client related (small fire) and non-client related

Cluster 4: Ambient – not Client related

Client Benefits:

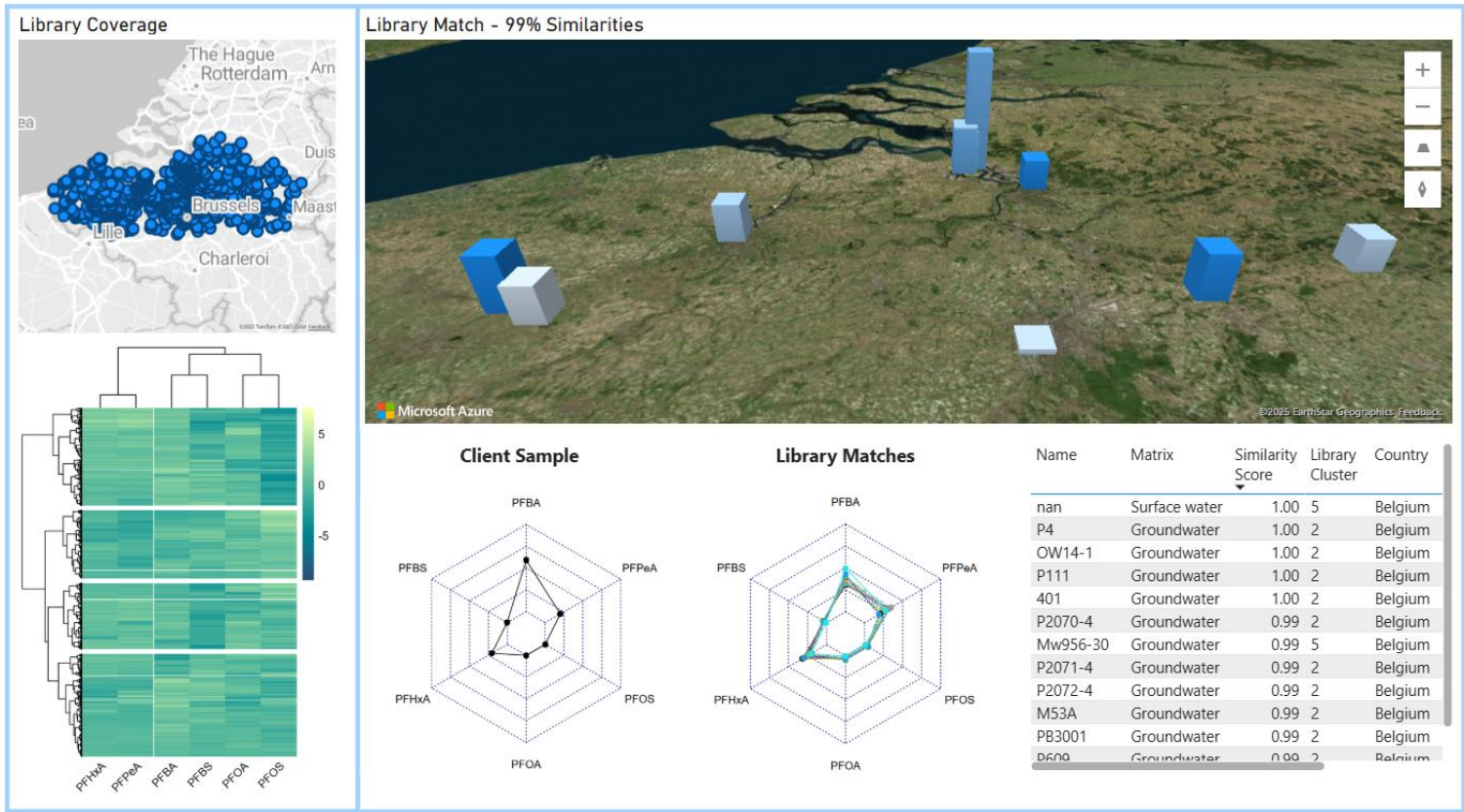
- Improved communication and understanding of PFAS data
- Reduced liability
- Cost reduction for future remediation and investigation



External Reference Data and Library Matching

Pattern Recognition is relative to the samples included in the analysis. If we want to examine the results in the broader context of PFAS occurrence at other sites and other source types, we must include reference library searches.

In this library search example, a project sample in Belgium showed a unique signature in a perimeter sample. The resulting library search of the signature returned several highly similar signatures in the publicly available Belgian dataset. This provides additional context and lines of evidence for potential offsite sources.

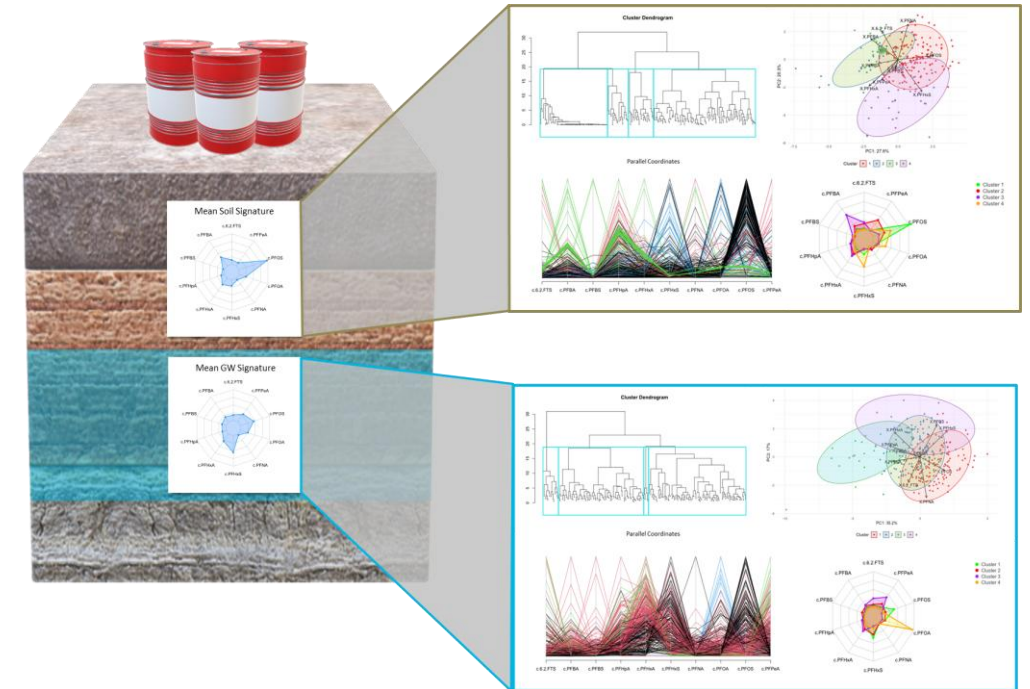


Profile Building – Project-Specific Libraries

We have a growing library of PFAS data that can be used to predict labels of new samples, including their source type and facility type. Exploratory methods let us find specific samples in the library that may match ours.

Project Example – Library Search for Anomalous Data

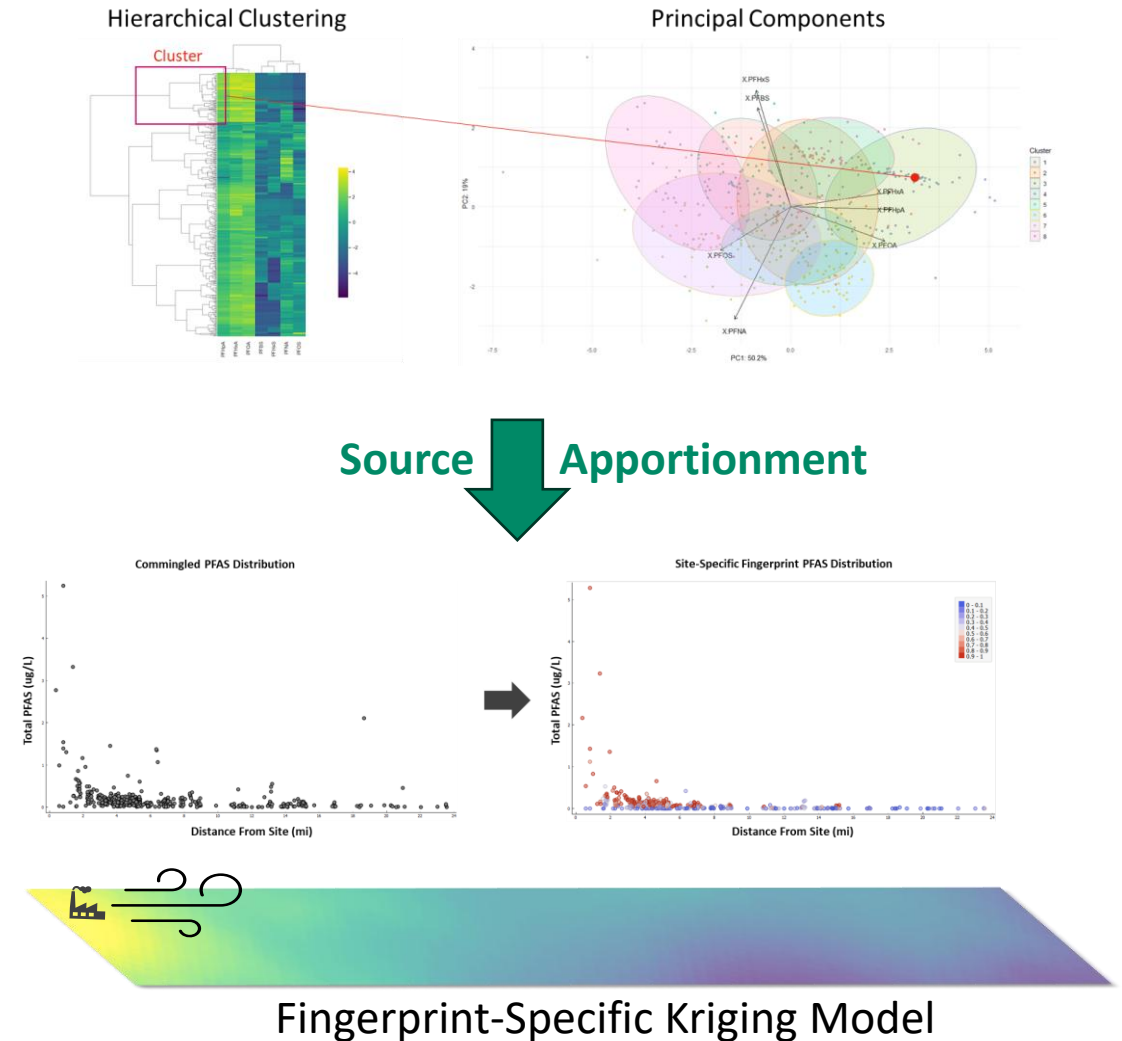
- Objective: Identify if anomalous sample in a site's PFAS pattern dataset is consistent with established program-wide profiles
- Methods: Clustering and Machine Learning Classification
- Outcome: Identified that the site data is highly similar to a subset of samples in the program library



Advanced Statistics for Source Isolation and Prediction

PFAS Profiles can be used to guide source apportionment and improve performance

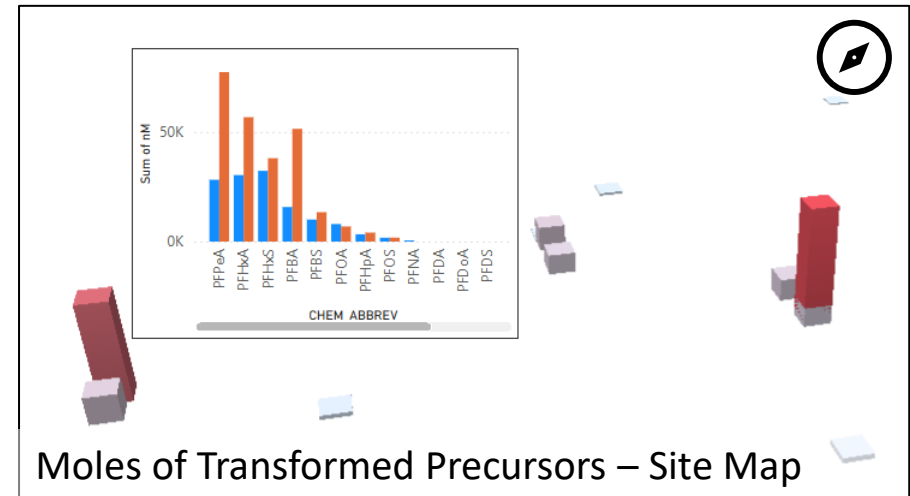
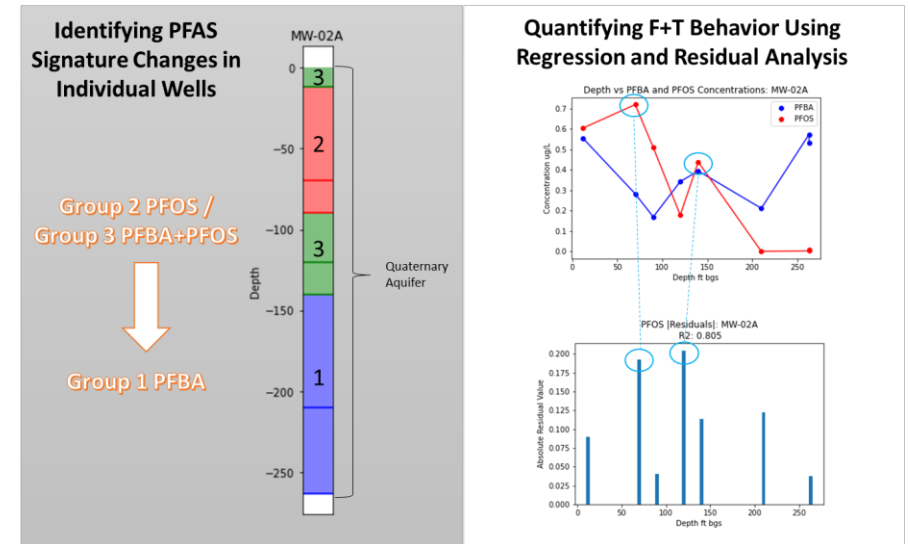
- Question – Can we predict the extent of groundwater PFAS concentrations contributed from a specific atmospheric source using only the existing onsite and offsite groundwater PFAS dataset?
- Outcome – Successfully modeled PFAS groundwater concentrations associated with the atmospheric source and identified several other potential sources in the area.



Incorporating Advanced Analytical Data

PFAS Profiles can appear to represent different sources, when in fact they might just represent transformation or sorption preferences associated with fate and transport.

- Question – How do we identify fate and transport behavior and differentiate its effects from our understanding of sources?
- Outcome – We can explore and account for these patterns in our analysis. We can also focus our efforts to collect specialized forensic samples from representative groups of samples, significantly reducing further investigation costs



Decision Framework – Adaptive Investigation Matrix (AIM)

- AIM provides a structured approach to the potential paths of a forensic investigation
- Key factors are flexible and can vary by client. They usually include:
 - Technical Feasibility
 - Outcome Certainty
 - Precedent/Regulatory Acceptance
 - Timeline
 - Cost

Example First-Phase Matrix

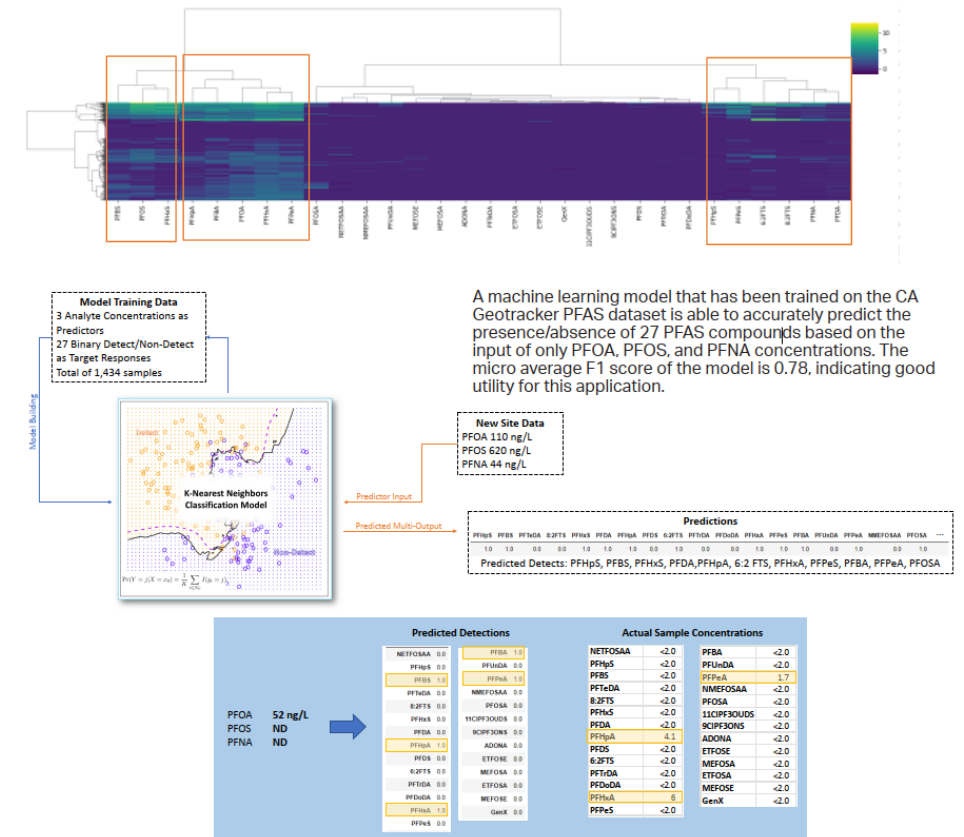
Strategy Recommendation	Scoring Metrics (Scored 1-5, Score of 5 equals most favorable to Client)			
	Certainty of Findings	Technical Feasibility	Implementation Time	Cost
Exploratory Pattern Screening via Clustering and PCA	4	5	5	5
Targeted Source Apportionment via PMF using Suspect Source Profiles	3	4	4	4
Advanced Branched Isomer Speciation	3	2	2	3
Selective and Focused Non-Targeted Analysis	3	3	2	1
Forensic Site History Evaluation	3	5	4	4
Atmospheric Emission Location(s) Spatial Analysis	4	3	3	4
Leapfrog Modeling Update to Include Forensic Indicators and Atmospheric Deposition Considerations	4	3	3	3

Conclusion and Future Directions

- Pattern screening is a chemometrics tool that allows us to understand our data more holistically and form a data-driven roadmap to incrementally narrow our investigation focus and efforts
- Can identify the most efficient options and tailor to any available data, small or large
- Part of a growing suite of dynamic chemistry tools we have at our fingertips
- Future directions include pairing chemometrics with PFAS fate and transport studies at AECOM's CleanTech Hub Laboratory with our in-house PFAS analysis capabilities

Future-Proofing our Pattern Screening

Predict Unmeasured PFAS Presence from Limited Analyte Lists



Thank You!

Please reach out to us with follow up questions:

Zachary Neigh Zachary.Neigh@aecom.com

William Leys William.Leys@aecom.com

Mathieu Charrier Mathieu.Charrier@aecom.com



Visit our PFAS website