Poster Reprint

ASMS 2026
ThP 357

Evaluation of a Consensus Based MS/MS Library Matching Algorithm and Integration into a Non-Targeted Software Platform for Increasing Exposome Coverage

David Schiessel[1], Jeremy P. Koelmel[1], Olivier Chevallier[2], David Godri[3], Randolph R. Singh[4], Thomas O. Metz[5], Emma E. Rennie[2], Cairo Ortiz[6], Parker Chang[1], Sheng Liu[1], and Krystal J. Godri Pollitt[1]

[1]Yale University,

[2]Agilent Technologies

[3]3rd Floor Solutions

[4]Columbia University

[5]PNNL

[6]Innovative Omics

Introduction

Motivating Improved PFAS Identification

PFAS identification workflows often rely on MS1-based metrics such as mass defect and homologous series behavior. While software like FluoroMatch includes advanced tools like Kaufmann plots to define PFAS chemical space (mass defect/carbon and mass/carbon domains), non-PFAS compounds may still fall within this region, leading to false positives. Incorporating MS/MS spectral matching provides orthogonal evidence to improve annotation confidence. In this work, we add MS/MS library searching and multi-metric scoring to FluoroMatch and other platforms to improve feature identification.

MS/MS similarity metrics each have limitations and may perform differently depending on spectral quality. Dot-product scoring is sensitive to noise, reverse cosine emphasizes shared fragments, and entropy similarity accounts for spectral complexity. These differences can lead to inconsistent rankings when used individually. Here, we implement a consensus-based scoring approach that combines multiple metrics to improve robustness and interpretation.

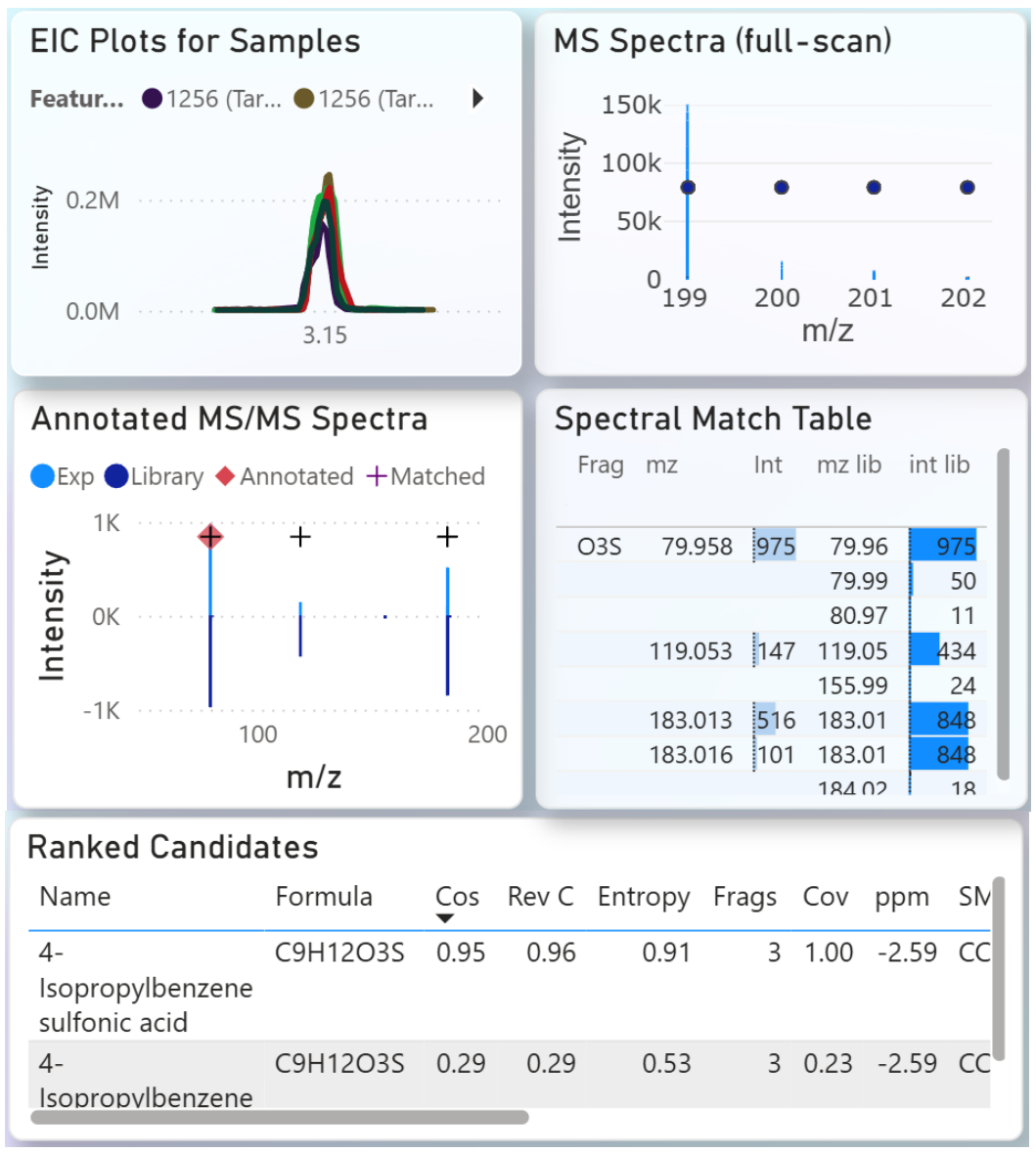

Figure 1: Visualization of MS2 Library Matches in FluoroMatch: High-confidence match for 4-Isopropylbenzenesulfonic acid in construction and landfill leachate. Matched against: MSBNK-EPA-ENTACT_AGILENT000009

Experimental

Reference Databases

MS/MS library matching was implemented within FluoroMatch to enable feature-level spectral annotation. For each detected feature, experimental MS/MS spectra were compared against candidate spectra from MassBank (v2025.10) and the NIST DIMSPEC Library. The DIMSpec library was used in its native format, while MassBank records were restructured into a unified SQLite database to enable efficient querying and scalable integration within the workflow.

Scoring Procedure

Candidate spectra were filtered by precursor m/z tolerance and ionization polarity prior to comparison. Experimental and reference spectra were aligned and filtered to remove low-quality signals before scoring. Spectral similarity was evaluated using three complementary metrics: dot-product, reverse cosine similarity, and entropy similarity. Candidates were ranked using a consensus-based approach that aggregates these scores.

Output

Users can filter (Figure 2) and prioritize identifications based on individual scoring metrics or combined ranking, depending on spectral quality and analysis goals. Visualization of the spectral match is also implemented (Figure 1).

Libraries: DIMSpec (native) | MassBank → SQLite
Filter: Precursor m/z (ppm) + polarity
Score: Dot-product | Reverse cosine | MSEntropy
Rank: Consensus scoring
User control: Filter candidates by individual metrics

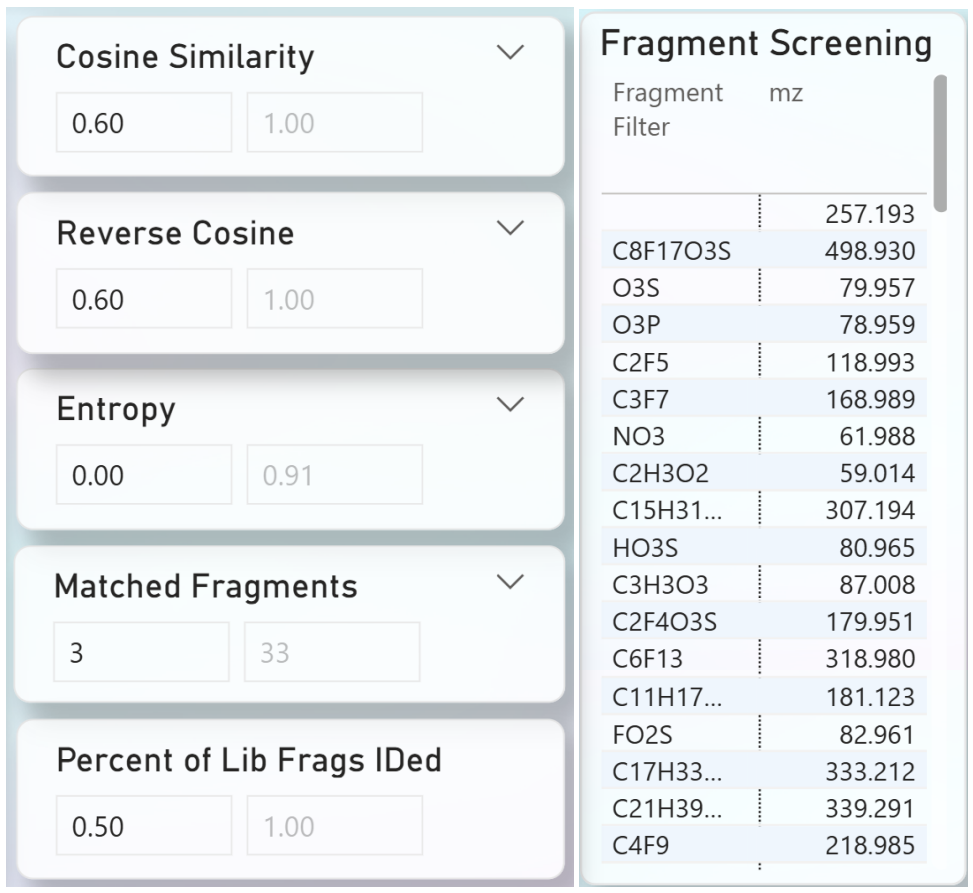

Figure 2: Visualizer allows user to filter features by scores and fragment occurrences

Identifying Limitations of MS1-Based Classification

MS1-based filtering using mass defect and homologous series resulted in features being assigned to PFAS-designated chemical space that were subsequently identified as non-PFAS. Features that aligned with expected mass defect/carbon and mass/carbon relationships, particularly when appearing as part of a homologous pattern, were initially flagged as PFAS candidates. These observations highlight a limitation of MS1-only approaches, where structurally unrelated compounds can satisfy PFAS classification criteria.

Resolving Misclassification with MS/MS Evidence

Incorporation of MS/MS spectral matching provided critical orthogonal evidence to refine this assignment. Library matching results did not support PFAS-like fragmentation patterns and instead indicated stronger similarity to non-PFAS reference spectra, enabling correction of the tentative identification. This demonstrates the value of integrating MS/MS evidence to reduce false positives and improve annotation confidence.

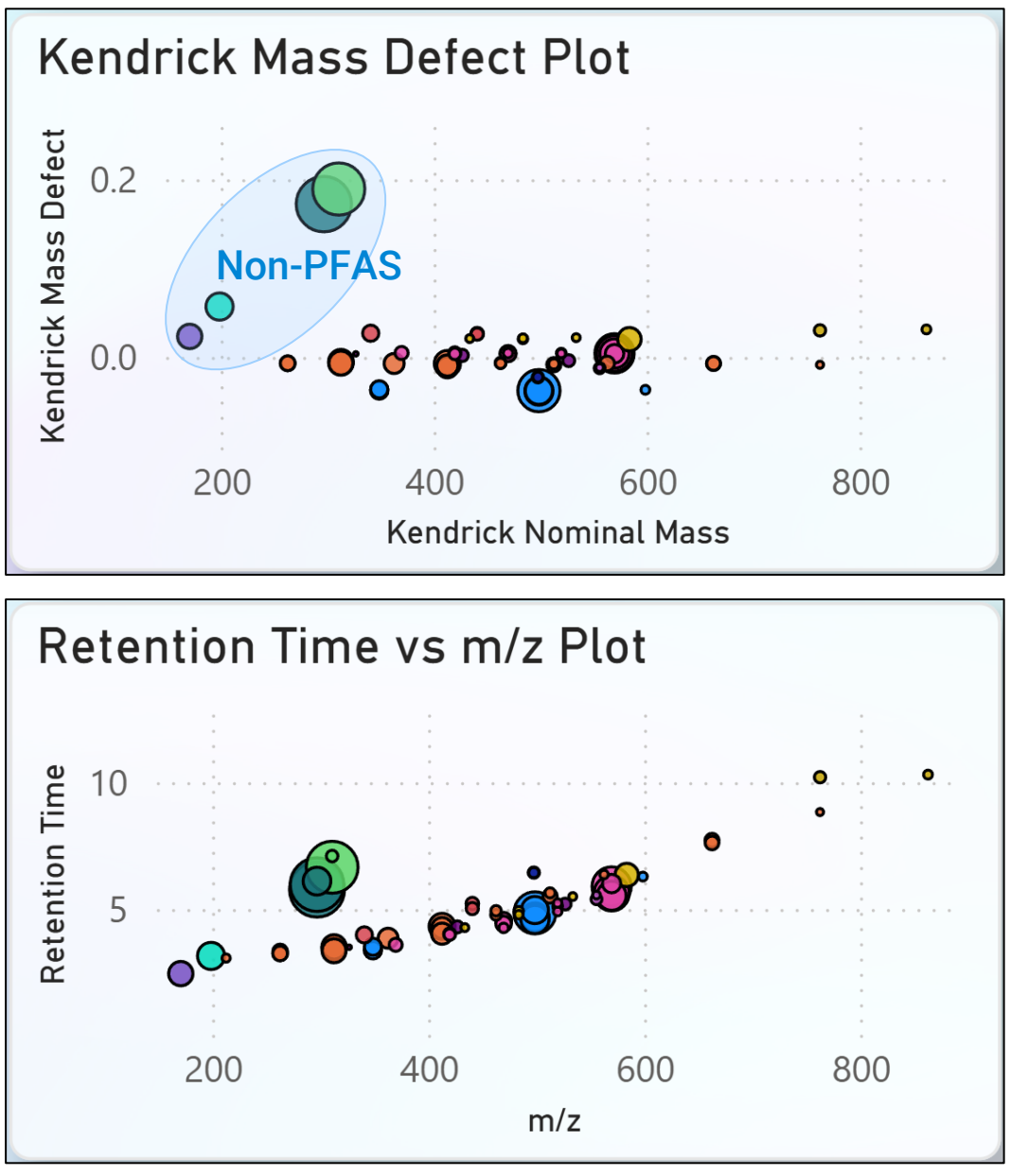

Figure 3: Homologous series of identified compounds showing PFAS and non-PFAS features

Improving Robustness with Consensus Scoring

Evaluation of similarity metrics further emphasized the importance of a multi-metric approach. In cases of noisy spectra, individual scores produced divergent results (e.g., low dot-product with high reverse cosine and entropy similarity), reflecting differences in sensitivity to intensity distribution and fragment matching. These observations highlight that no single metric is sufficient across all conditions and support the use of a consensus-based scoring framework.

Example: Resolving Misclassification

Isopropylbenzenesulfonic acid (C_9, m/z = 199.0434) was initially classified within PFAS chemical space based on MS1 characteristics, including its mass defect and position in the Kaufmann mass defect/carbon and mass/carbon domains (Figure 4). The presence of a homologous feature further supported a tentative PFAS assignment. However, MS/MS library matching did not support PFAS-like fragmentation and instead indicated stronger similarity to non-PFAS reference spectra. This example highlights how MS1-based metrics alone can lead to false positives and demonstrates the importance of incorporating MS/MS evidence to refine feature identification

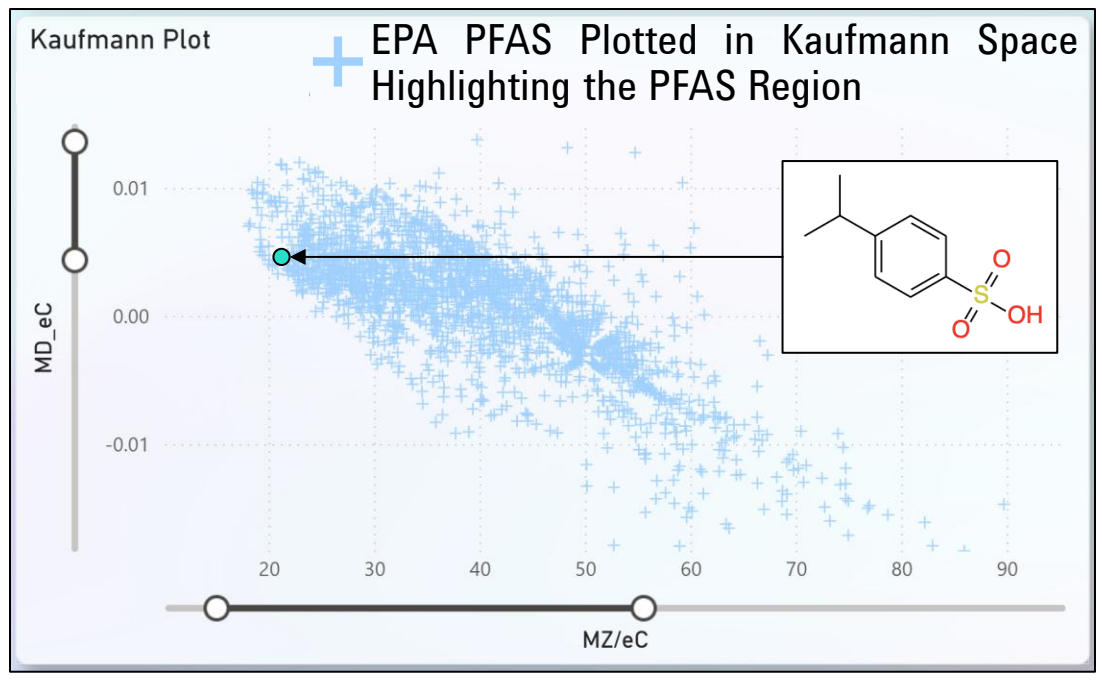

Figure 4: Example of non-PFAS within Kaufmann Chemical Space (4-Isopropylbenzenesulfonic acid).

Results and Discussion

Application: Identifying Co-Contaminants in Construction Landfill Samples

In addition to 68 Annotated PFAS features (Figure 3), MS/MS library matching enabled identification of several non-PFAS compounds in construction and demolition landfill samples. These included phenolic antioxidants (e.g., Fenozan, 4,4'-Thiobis(6-tert-butyl-m-cresol)), organophosphate additives (e.g., tributyl phosphate), and linear alkylbenzene sulfonates (LAS). These compounds are commonly associated with materials found in construction debris, including plastics, insulation, coatings, and industrial additives (Figure 5).

Linear alkylbenzene sulfonates are widely used as surfactants in detergents and cleaning agents and may enter landfill environments through residual materials or wastewater-impacted debris. Phenolic antioxidants are used as stabilizers in polymers and rubber, while organophosphate esters serve as plasticizers and flame retardants. Their detection highlights the complex mixture of anthropogenic chemicals present in landfills and demonstrates the value of expanding analysis beyond PFAS.

These findings illustrate that incorporation of broad MS/MS libraries (e.g., MassBank) enables identification of co-contaminants and provides a more comprehensive view of chemical exposures in environmental samples.

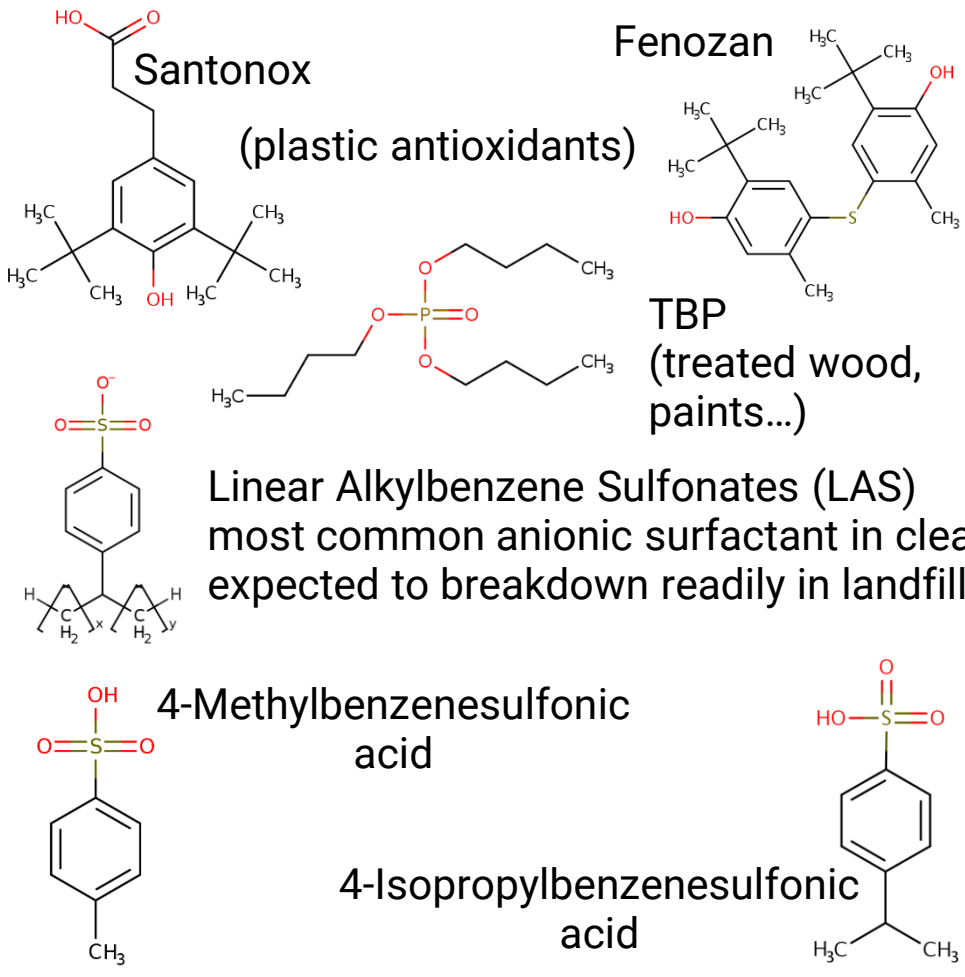

Figure 5: Co-contaminant features in leachate

Conclusions

Consensus MS/MS Scoring Improves Identification and Expands Coverage

- Implemented MS/MS library matching and consensus-based scoring within FluoroMatch
- Demonstrated that MS1-based PFAS classification can produce false positives, particularly for compounds occupying Kaufmann chemical space
- Integration of MS/MS evidence enables correction of misassignments and improves identification confidence
- Multi-metric scoring (dot-product, reverse cosine, entropy) provides robust evaluation across variable spectral quality, but requires careful interpretation
- Incorporation of MassBank expands identification beyond PFAS to include co-contaminants, increasing exposome coverage

References

[1]MassBank (v2025.10), Zenodo

[2] NIST DIMSpec Library, J. Am. Soc. Mass Spectrom., 2024

[3] Spectral entropy outperforms MS/MS dot product similarity for small-molecule compound identification, Nature Methods 18, 1524–1531 (2021)

Software Availability

Download the free open-source software:

InnovativeOmics.com

Contact

David@Schiessel.us

Acknowledgements

Nicole Robey, Keith Weitz, Thabet Tolaymat, and Kate Bronstein led the construction and demolition project and provided samples.

https://www.agilent.com/en/promotions/asms

DE-014516

Published in USA, June 30, ,2026

Agilent

Trusted Answers