

# A deconvolution and library search algorithm for comprehensive non-target analysis using LC-HR-MS data independent acquisition

Saer Samanipour, Malcolm J. Reid, and Kevin V. Thomas

16<sup>th</sup> International Conference on  
Chemistry and the Environment  
(Oslo June 18<sup>th</sup> to 22<sup>nd</sup>)



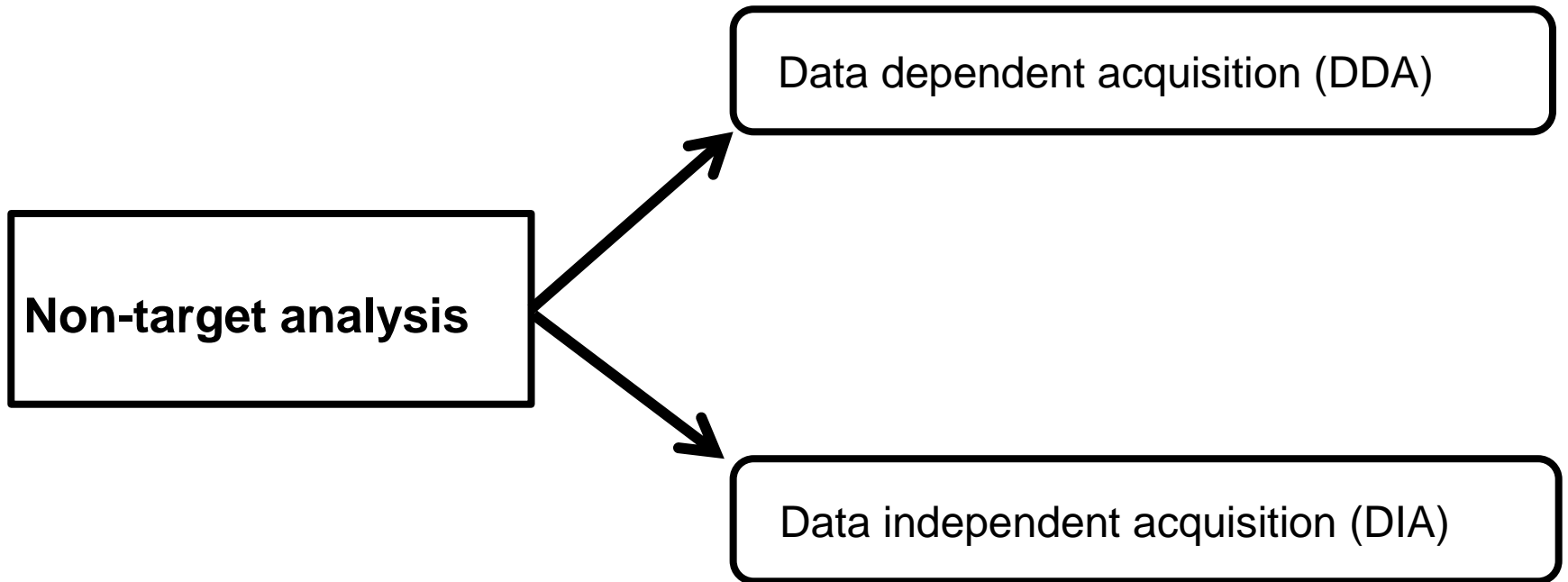
# Analysis overview



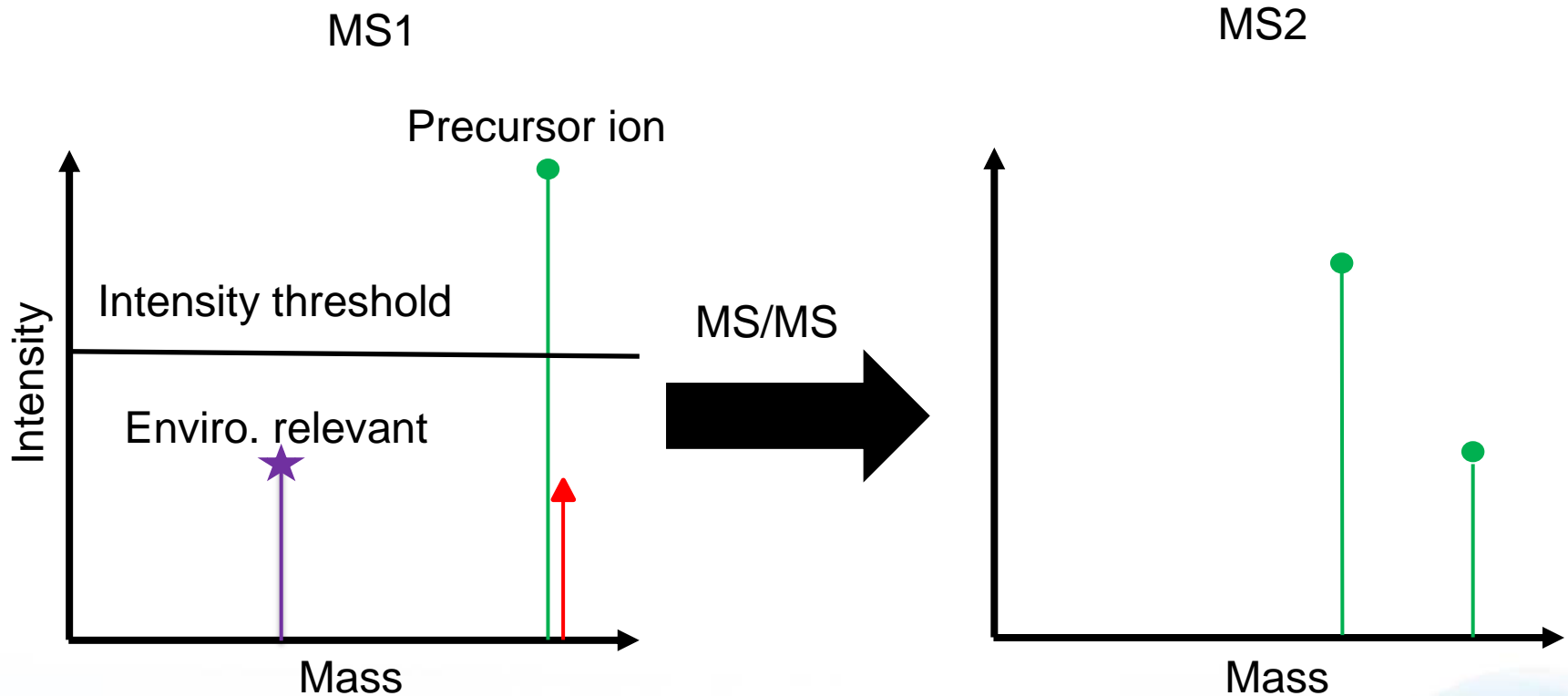
1. Around 90% of features are left for non-target analysis.
2. Mixture of naturally occurring and anthropogenic compounds.
3. We do not know about the environmental relevance of these features.

Non-target analysis is one of the most comprehensive approaches for analysis of complex environmental samples.

# Non-target analyses of complex samples

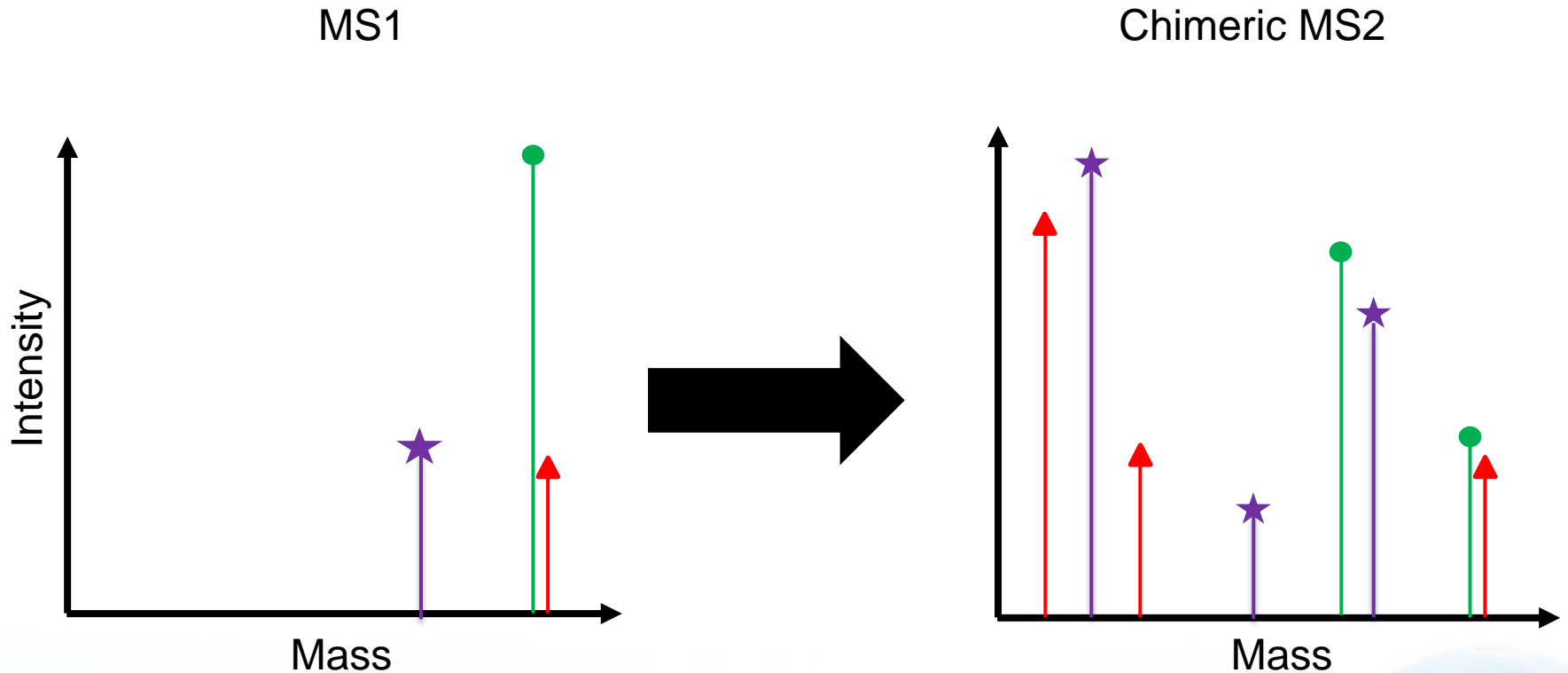


# Data dependent acquisition



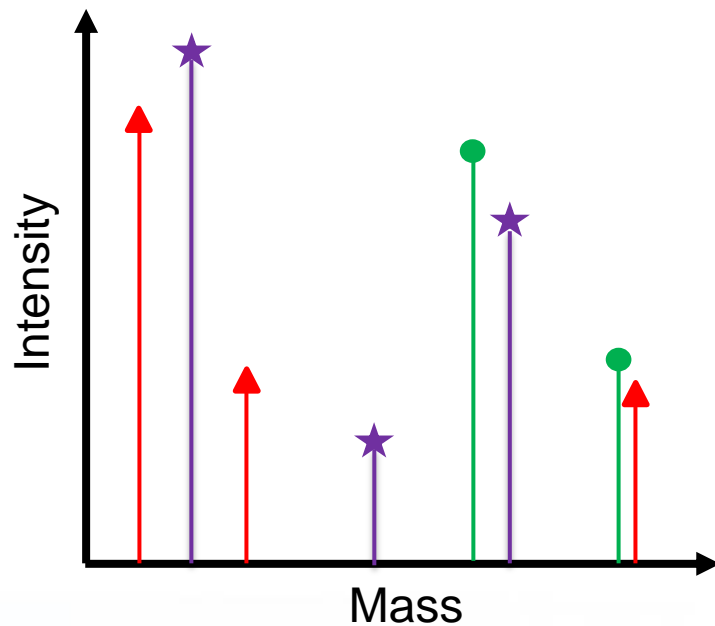
This method is highly intensity-dependent and not adequate for discovery based non-target analysis.

# Data independent acquisition



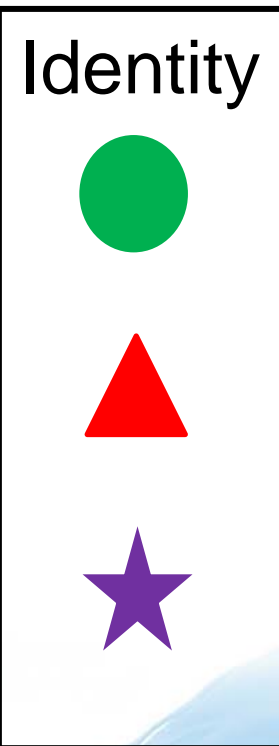
The DIA method generates highly complex chimeric spectra which require proper deconvolution and library search algorithms.

# Chimeric MS2 spectra via DIA



Deconvolution

+  
ULSA



# Semi-synthetic data generation

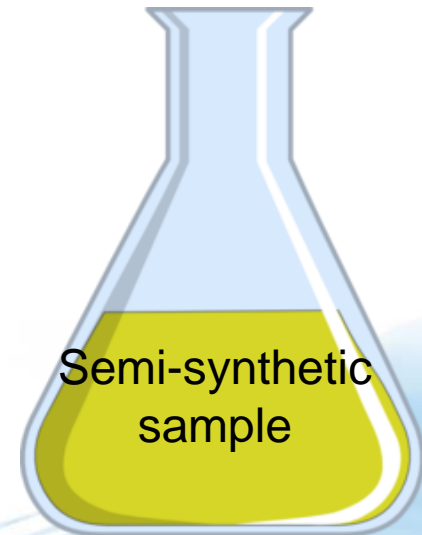
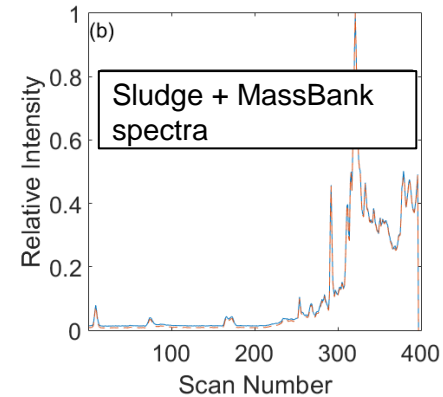
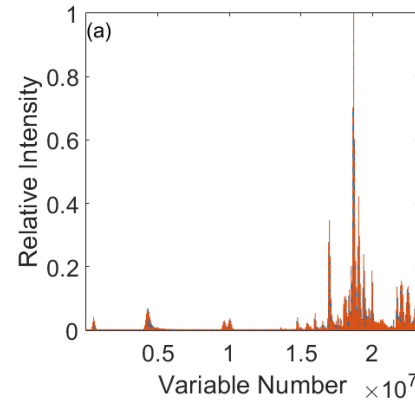
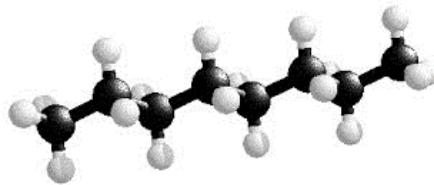
15 Sludge samples



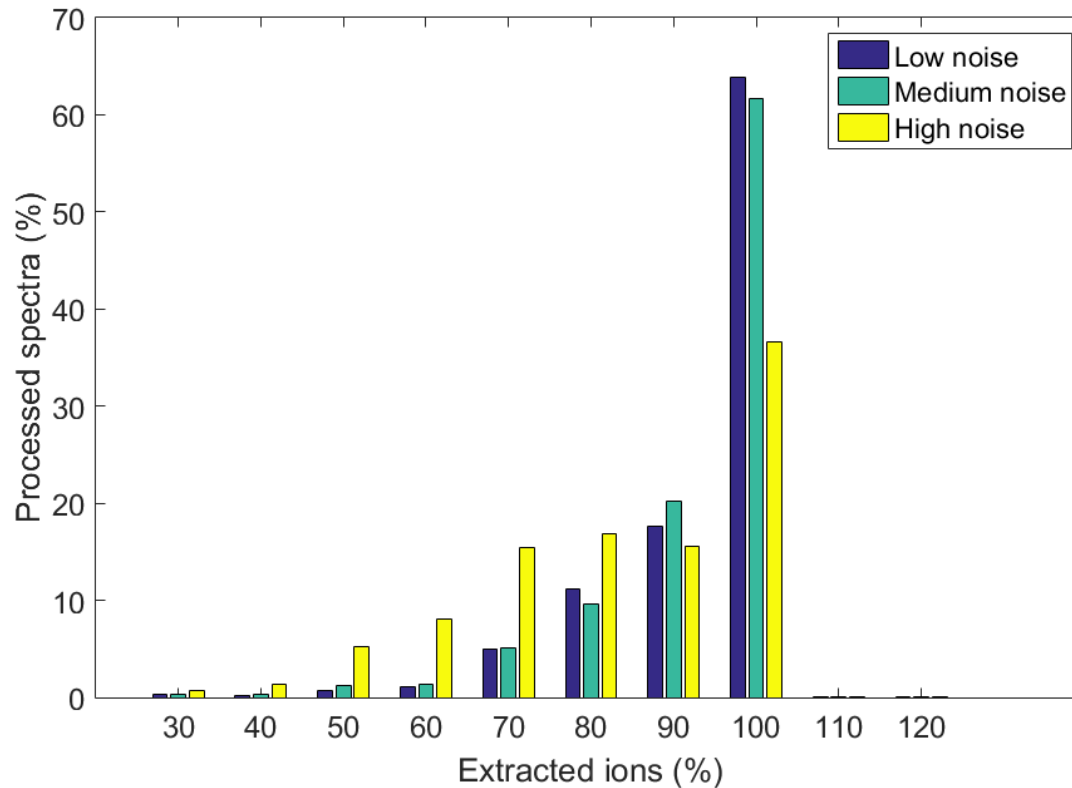
Artificial introduction



6000 MassBank



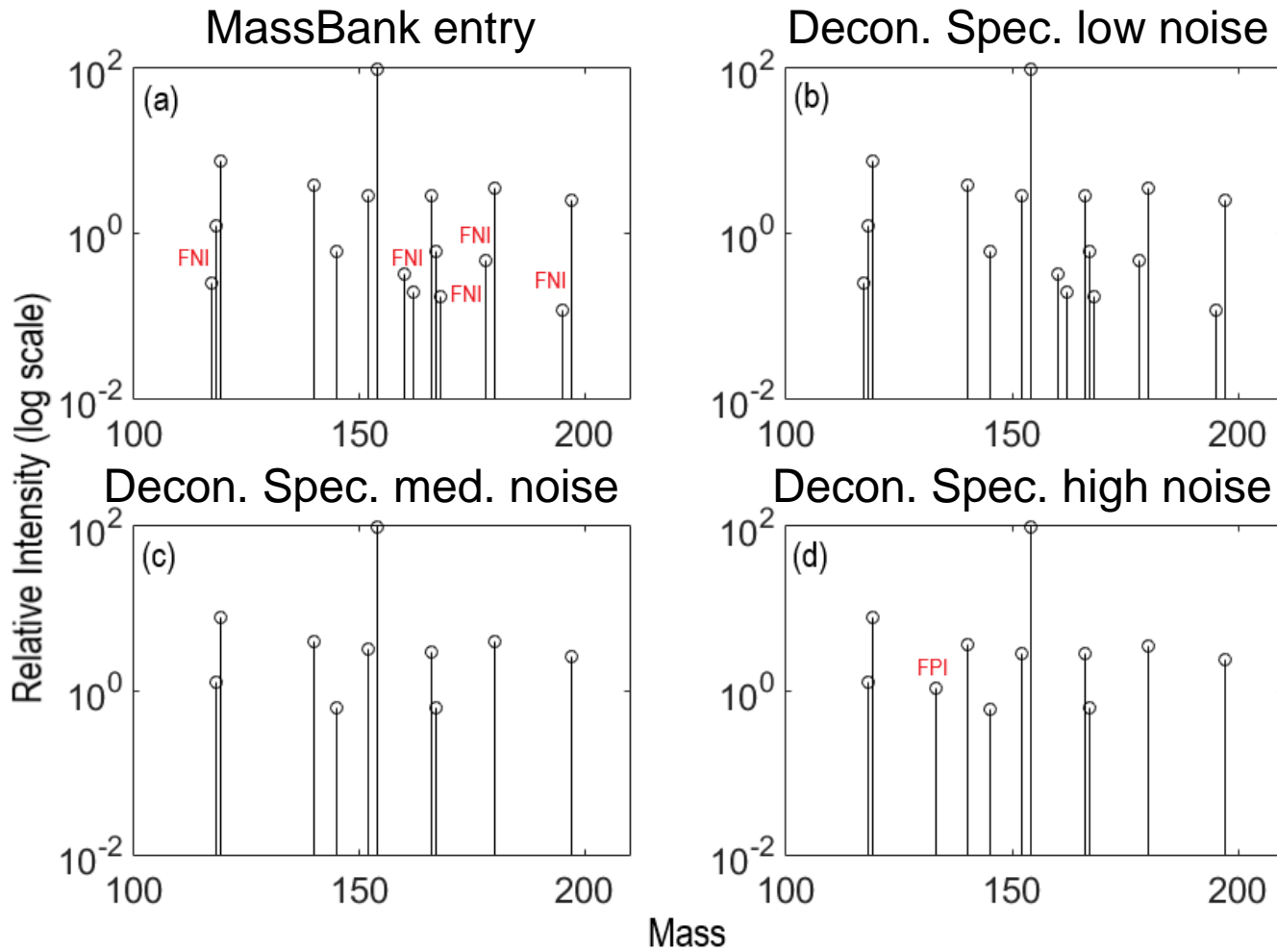
# Deconvolution algorithm



More than 60% of the added spectra were successfully extracted for all three levels of noise without peak picking or peak modeling.

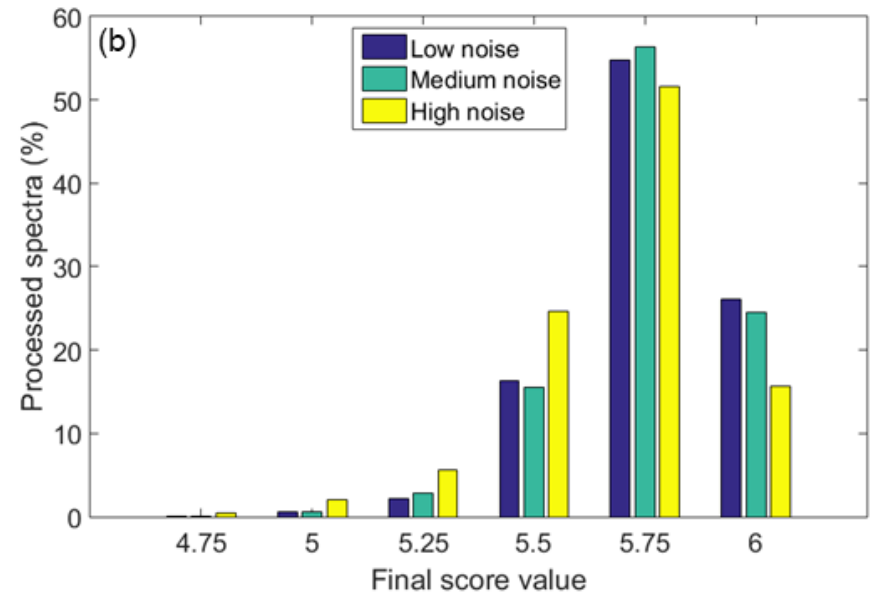
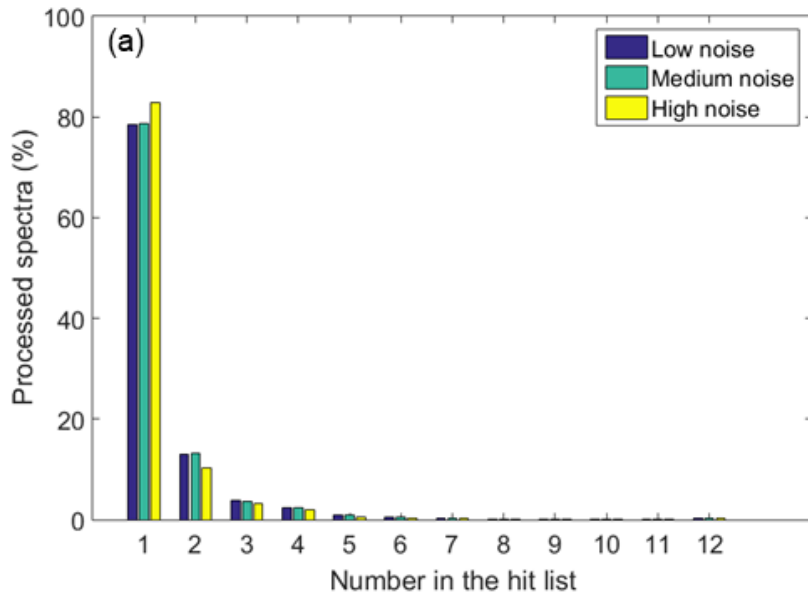


# Deconvolution algorithm



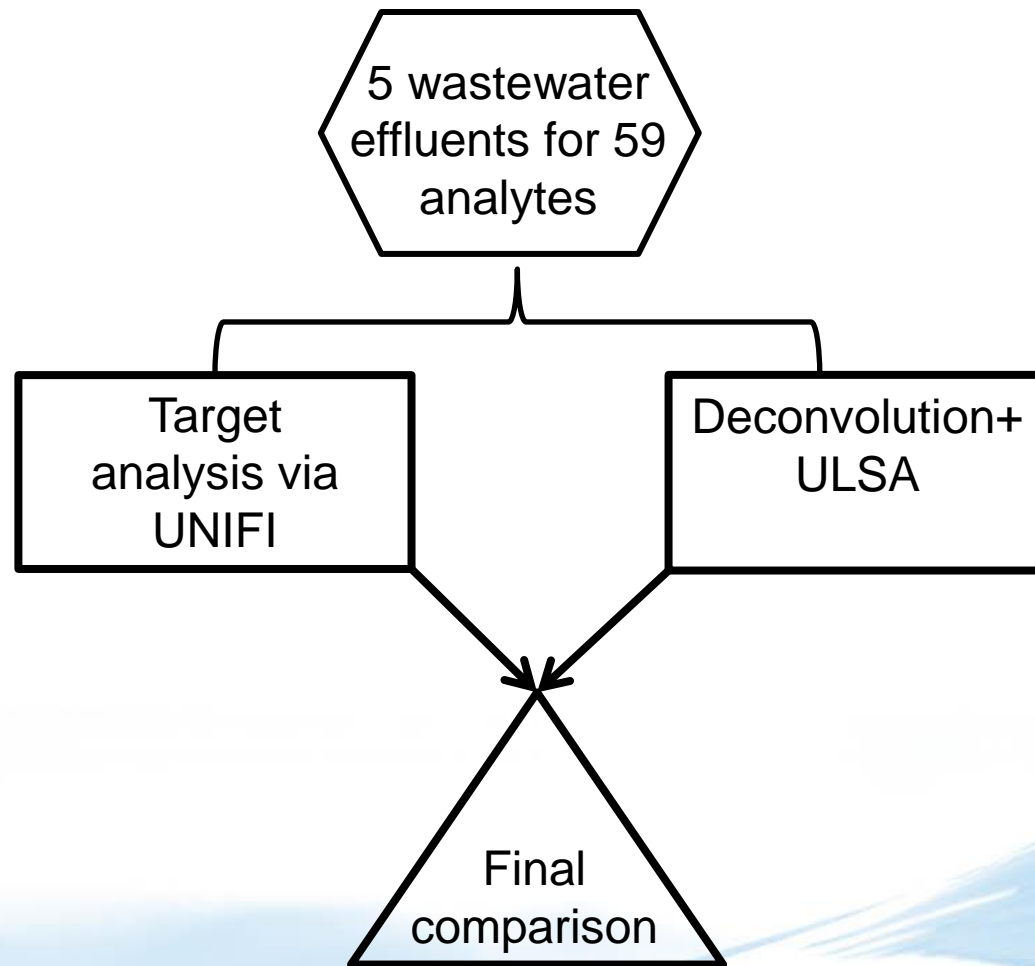
FNI stands for false negative ion whereas FPI stands for false positive ion.

# Universal library search algorithm



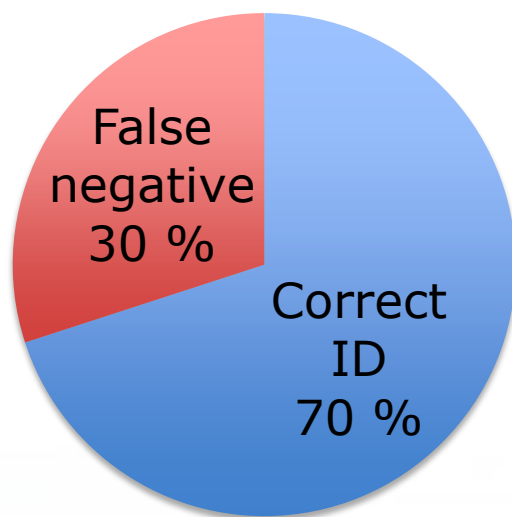
More than 95% of the added MassBank spectra ranked among the top three in the final hit list at all three levels of noise.

# Real environmental samples



# Real environmental samples

## Detection results via combination of Decon. algo. And ULSA



False  
positive  
0 %

All the false negative cases were caused by low intensity ions (i.e. smaller than 800 counts). Therefore the deconvolution algorithm was not able to extract the MS2 spectra.

# Conclusions

- The developed deconvolution algorithm was able to extract more than 60% of the added spectra for 95% of the processed spectra.
- This deconvolution algorithm does not rely on peak picking or peak modeling.
- The ULSA was able to rank the correct spectra among the top three hits in the final hit list for more than 95% of the processed spectra.
- The ULSA did not produce any cases of false positive detection for both the semi-synthetic data and the real environmental samples.
- The ULSA provides the highest level of cross-platform compatibility for a library search algorithm.
- The combination of these two algorithms can be integrated into the existing discovery-based non-target workflows.

# Thank you for your attention!

Contact info:

Saer Samanipour: [saer.samanipour@niva.no](mailto:saer.samanipour@niva.no)

Malcolm J. Reid: [malcolm.reid@niva.no](mailto:malcolm.reid@niva.no)



$$\begin{aligned}
S_f = & w_1 \cdot \frac{f_{\text{matched}}}{f_{\text{user}}} \\
& + w_2 \cdot \frac{f_{\text{matched}}}{f_{\text{reference}}} \\
& + w_3 \cdot \frac{MS_{\text{tol}} - MS_{\text{err,per}}}{MS_{\text{tol}}} \\
& + w_4 \cdot \frac{f_{\text{matched}}}{f_{\text{reference}}} \cdot \frac{MS_{\text{tol}} - |MS_{\text{err,frag}}|}{MS_{\text{tol}}} \\
& + w_5 \cdot \frac{f_{\text{matched}}}{f_{\text{reference}}} \cdot \frac{2 \cdot MS_{\text{tol}} - STDEV_{\text{err,frag}}}{2 \cdot MS_{\text{tol}}} \\
& + w_6 \cdot MF_{\text{forward}} \\
& + w_7 \cdot MF_{\text{reverse}}
\end{aligned}$$

Number of matched frag in the user spec

Number of matched frag in the ref spec

Mass error of the parent ion

Mass error of the matched fragments

Standard deviation of mass error

Forward match factor

Reverse match factor