machine learning TO SUPPORT QUANTITATIVE NTS

Anneli Kruve

kruvelab.com


it all boils down to
RISK

it all boils down to

RISK

Highly toxic substances deserve attention

TOXICITY

it all boils down to

RISK

Highly toxic and high exposure substances deserve attention

TOXICITY

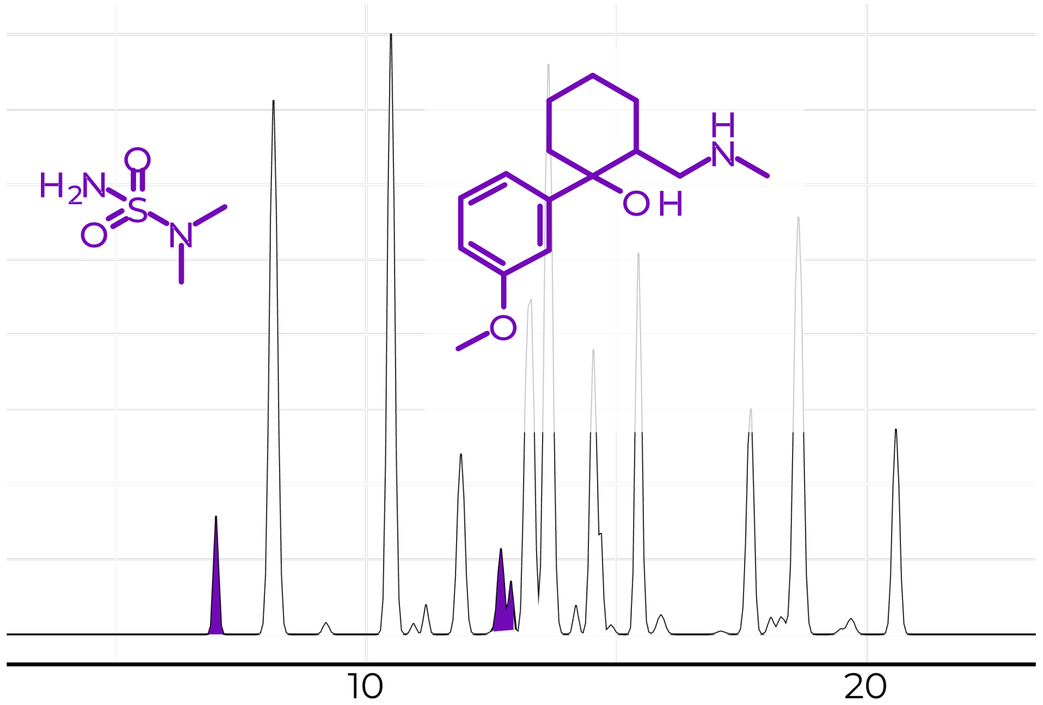
EXPOSURE

quantification

IN LC/HRMS

quantification

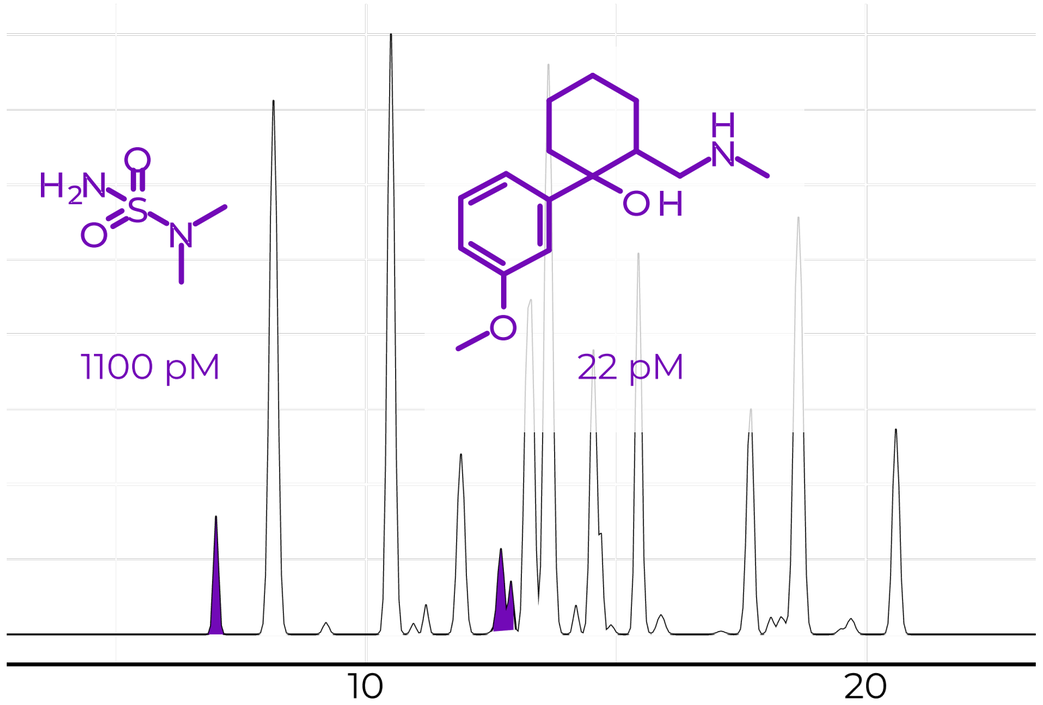
IN LC/HRMS



Malm et al. Molecules **2021**, *26* (12), 3524.

quantification

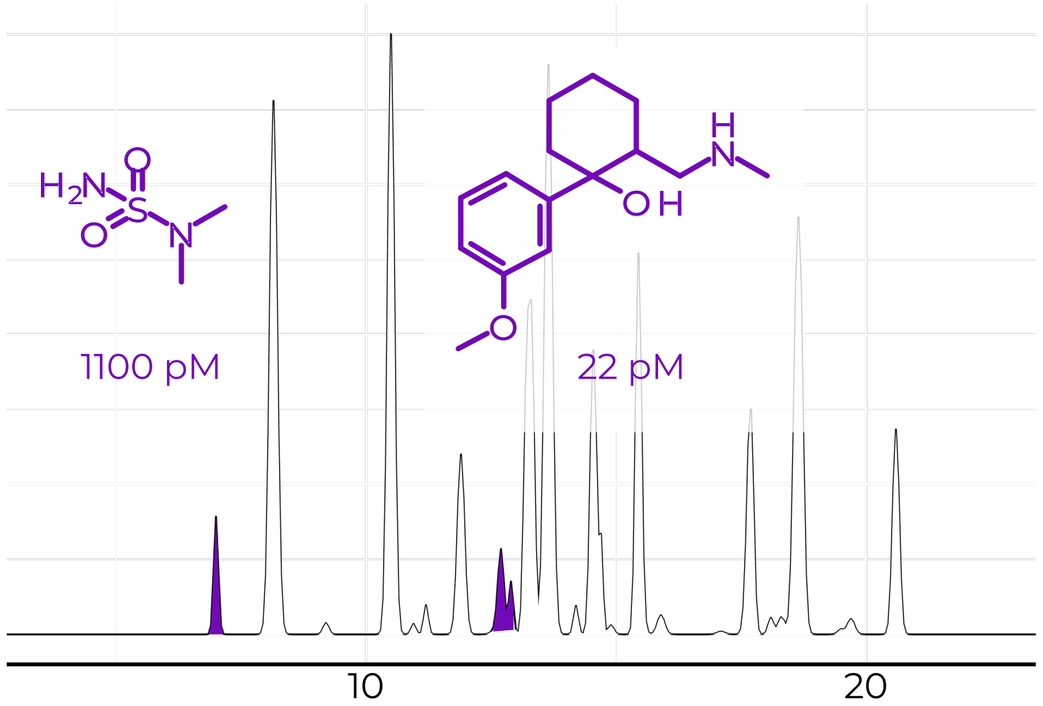
IN LC/HRMS



Malm et al. Molecules **2021**, *26* (12), 3524.

quantification

IN LC/HRMS

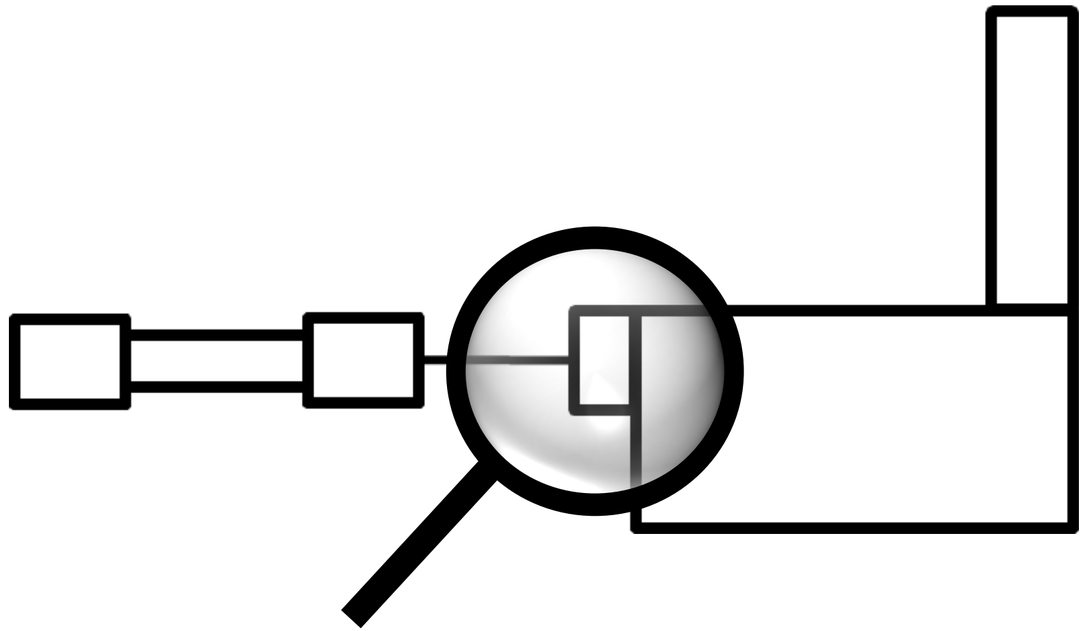


Response up to 10,000,00x different

Malm et al. Molecules **2021**, *26* (12), 3524.

quantification

IN LC/HRMS



ELECTROSPRAY

Malm et al. Molecules **2021**, *26* (12), 3524.

ionization is

IMPACTED

CHEMICAL STRUCTURE

MOBILE PHASE

MATRIX & INSTRUMENT

quantification

METHODS

quantification

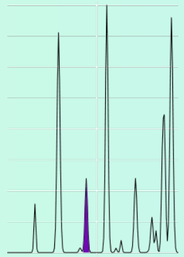
METHODS

Tested in NORMAN interlaboratory comparison

Malm et al. Anal. Chem. **2024**, *96* (41), 16215-16226.

quantification METHODS

Tested in NORMAN interlaboratory comparison

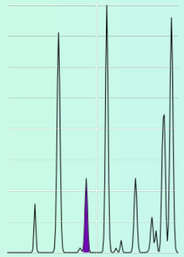


CLOSE ELUTING COMPOUNDS

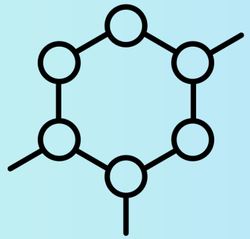
Malm et al. Anal. Chem. **2024**, *96* (41), 16215-16226.

quantification METHODS

Tested in NORMAN interlaboratory comparison



CLOSE ELUTING COMPOUNDS

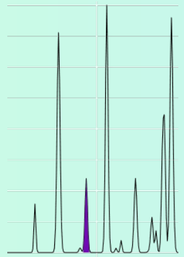


STRUCTURALLY SIMILAR COMPOUNDS

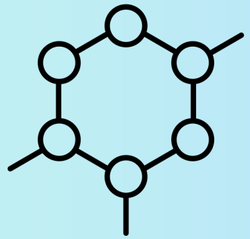
Malm et al. Anal. Chem. **2024**, *96* (41), 16215-16226.

quantification METHODS

Tested in NORMAN interlaboratory comparison



CLOSE ELUTING COMPOUNDS

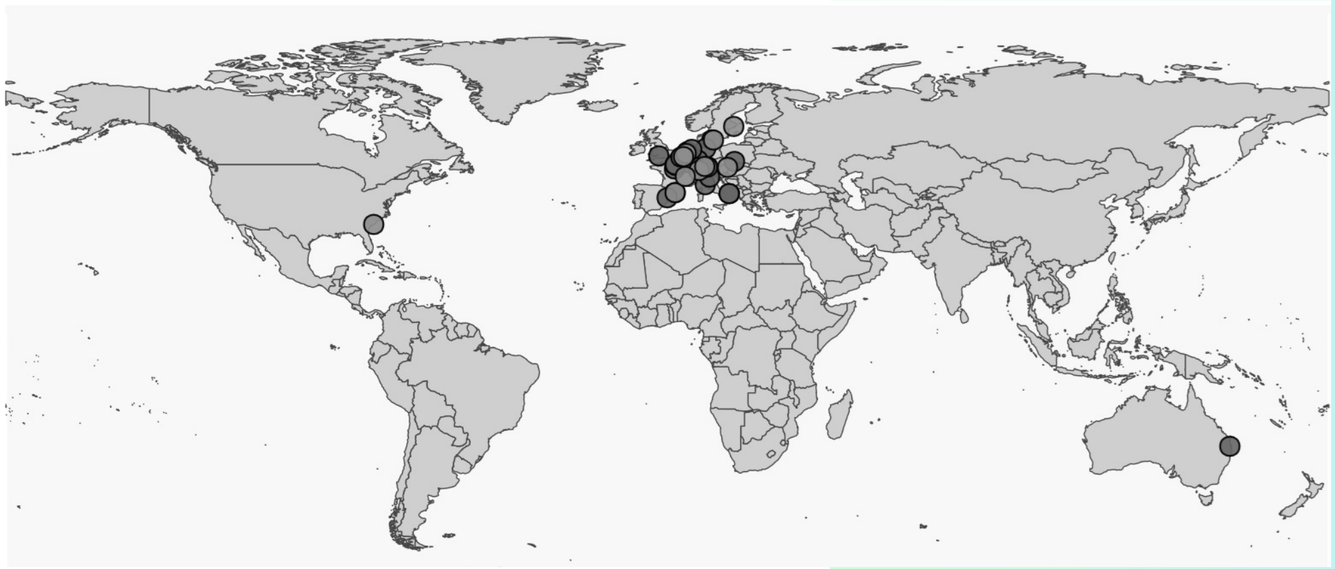


STRUCTURALLY SIMILAR COMPOUNDS

IE PREDICTIONS WITH ML

Malm et al. Anal. Chem. **2024**, *96* (41), 16215-16226.

interlaboratory COMPARISON



37/41
Labs/methods

45
Compounds

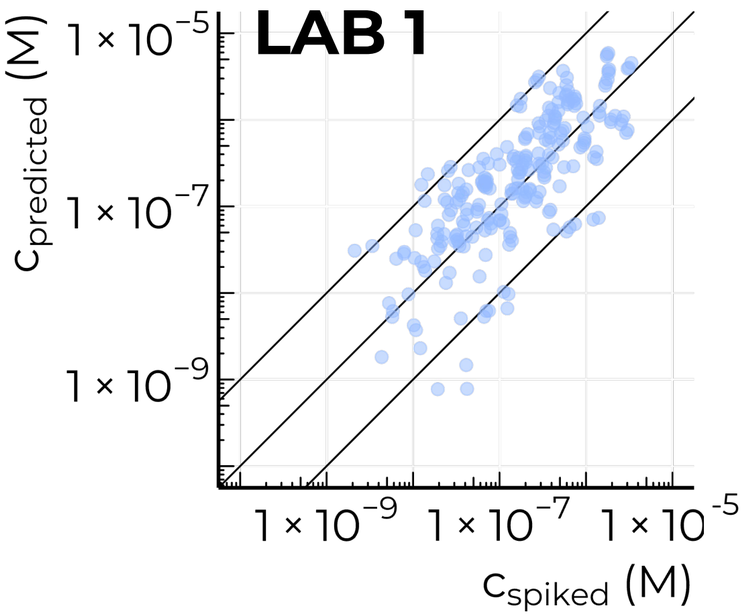
3
Sample types

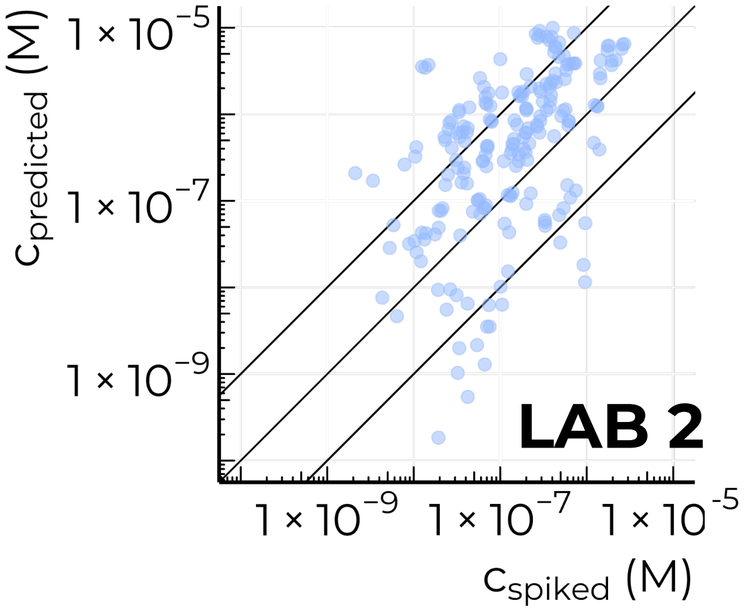
2
Concentration levels

Malm et al. Anal. Chem. **2024**, *96* (41), 16215-16226.

ACCURACY

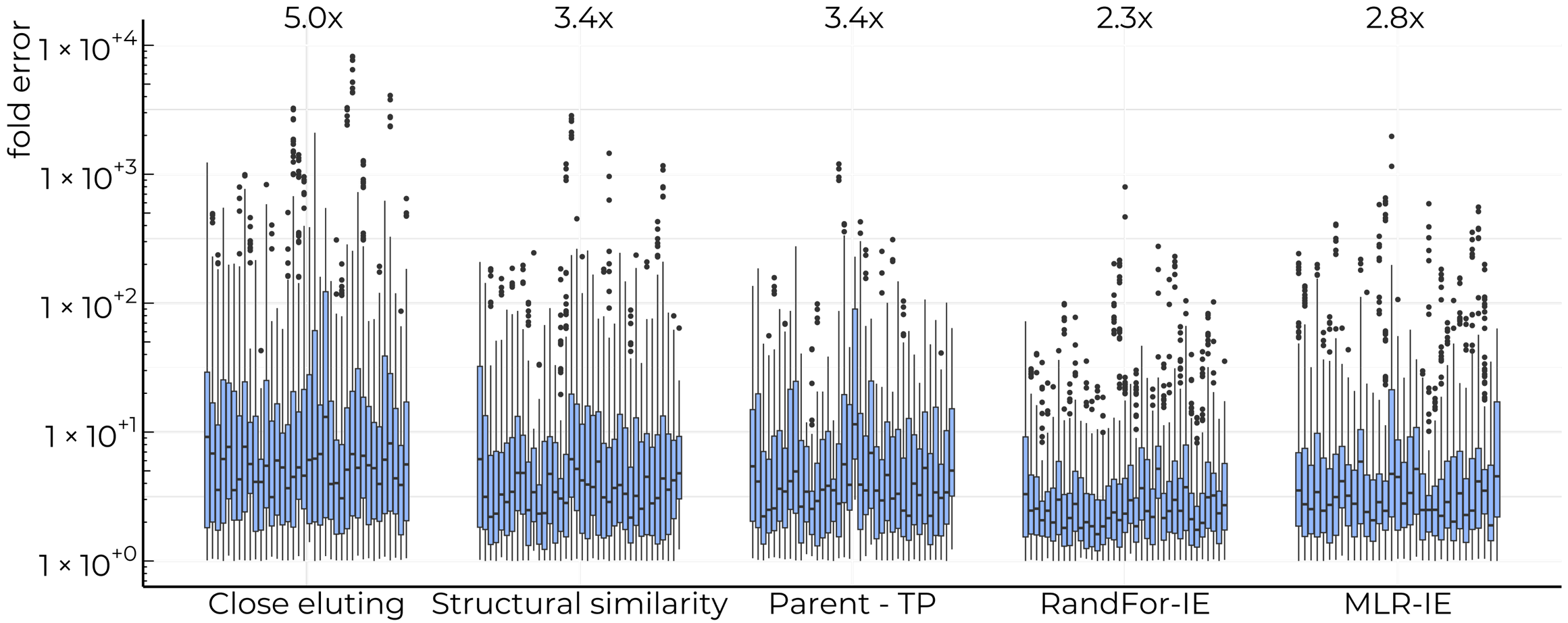
Predicted vs spiked concentration





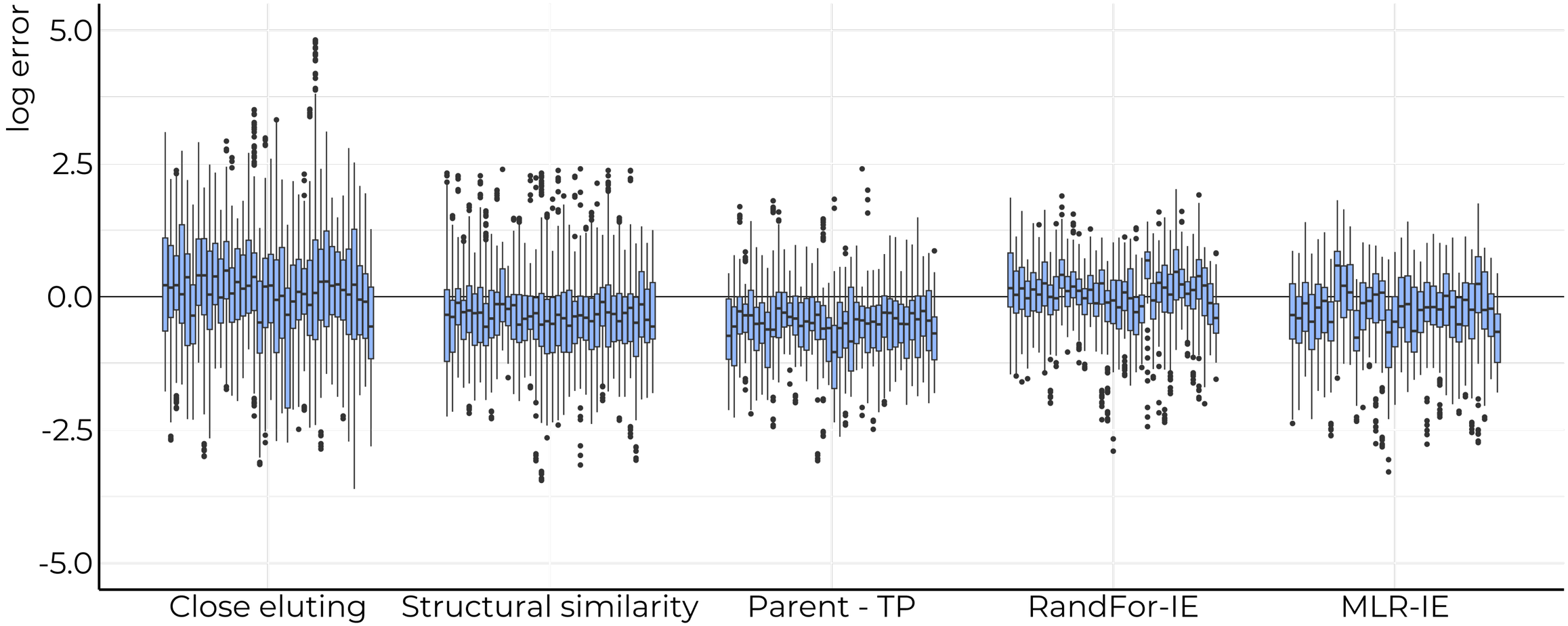
Malm et al. Anal. Chem. **2024**, *96* (41), 16215-16226.

FOLD ERROR



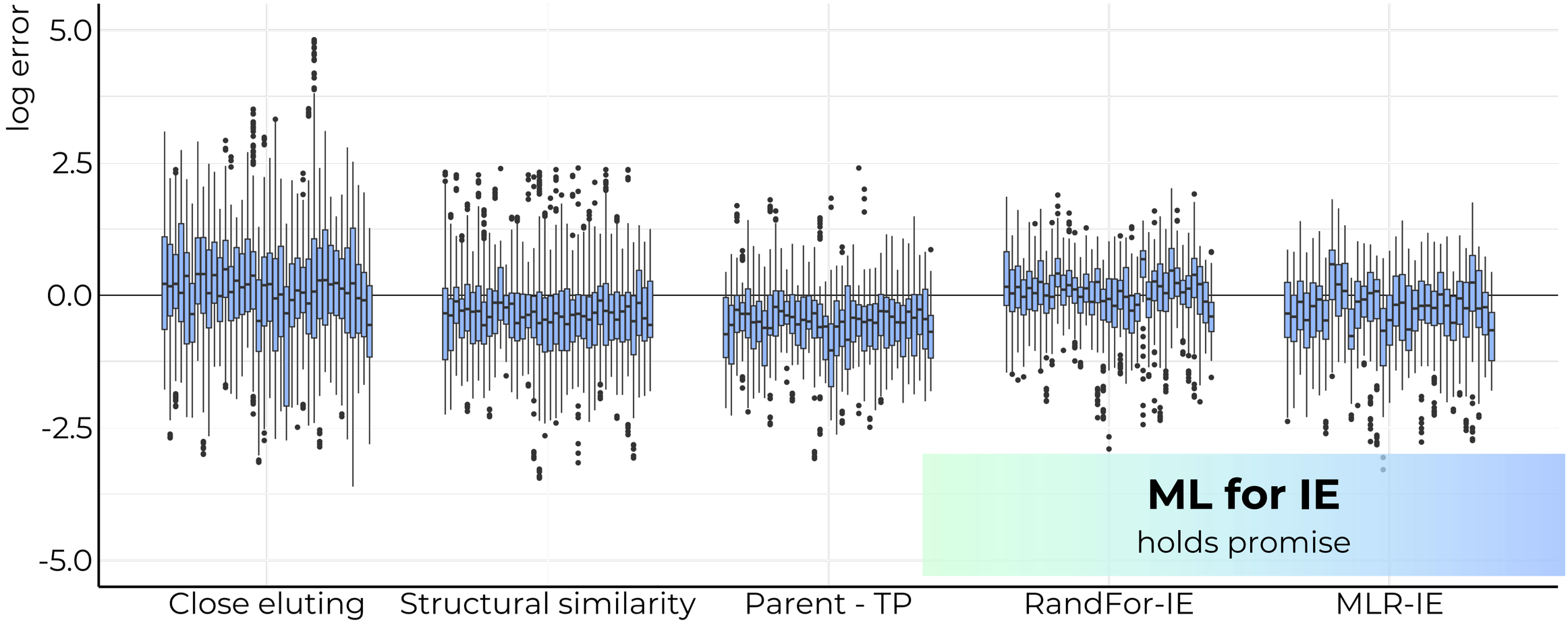
Malm et al. Anal. Chem. **2024**, *96* (41), 16215-16226.

LOG ERROR



Malm et al. Anal. Chem. **2024**, *96* (41), 16215-16226.

LOG ERROR



Malm et al. Anal. Chem. **2024**, *96* (41), 16215-16226.

future

OF QUANTIFICATION

open

QUESTIONS

Where should we go in next 3 years?

? **CHEMICAL SPACE**

? **UNIDENTIFIED CHEMICALS**

? **RISK**

open

QUESTIONS

Where should we go in next 3 years?

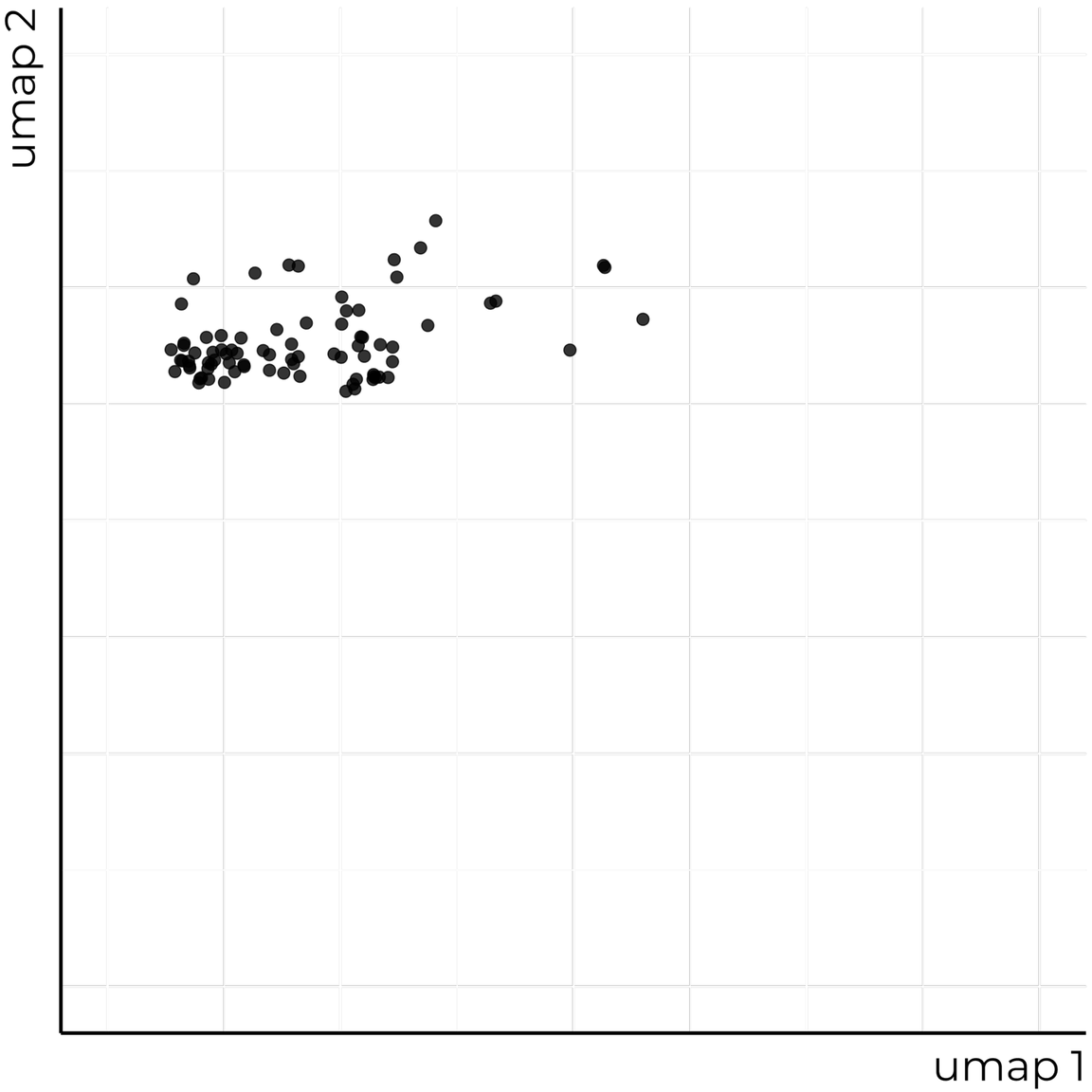
? CHEMICAL SPACE

? UNIDENTIFIED CHEMICALS

? RISK

PROBLEM

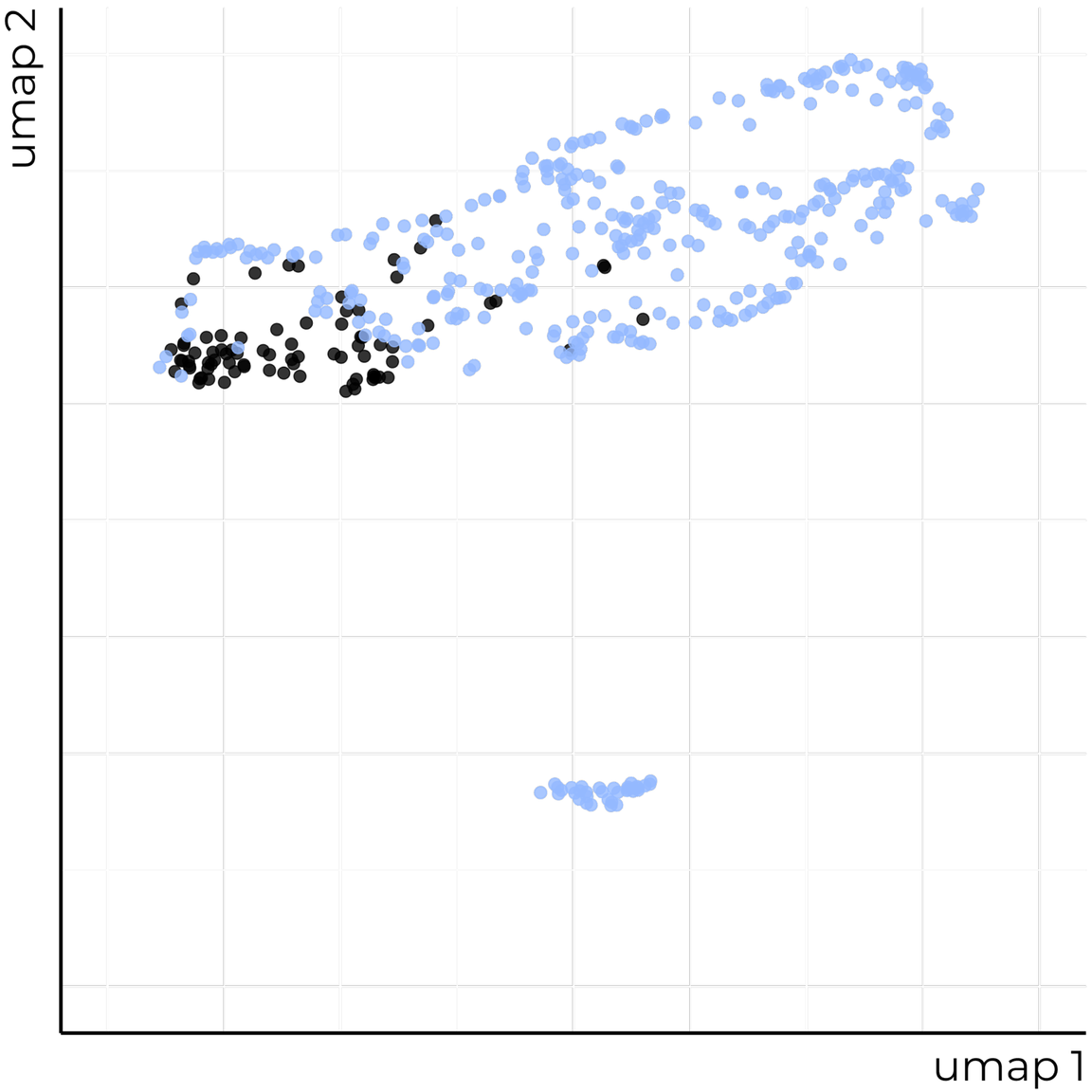
Chemical space coverage



Wang et al. in preparation.

PROBLEM

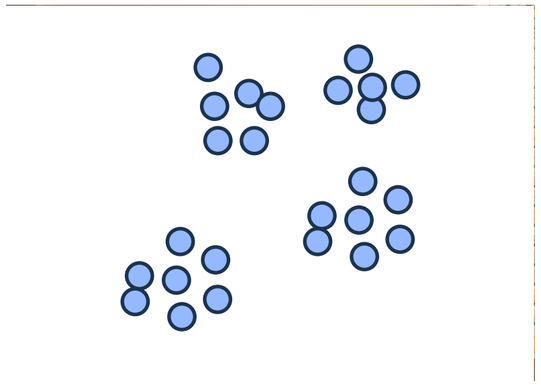
Chemical space coverage



Wang et al. in preparation.

chemical space

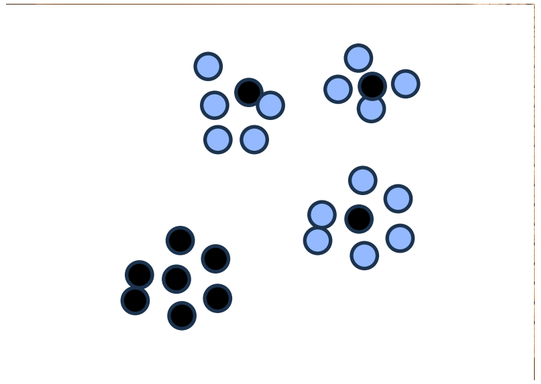
EXPLO-RATION



Wang et al. in preparation.

chemical space

EXPLO-RATION

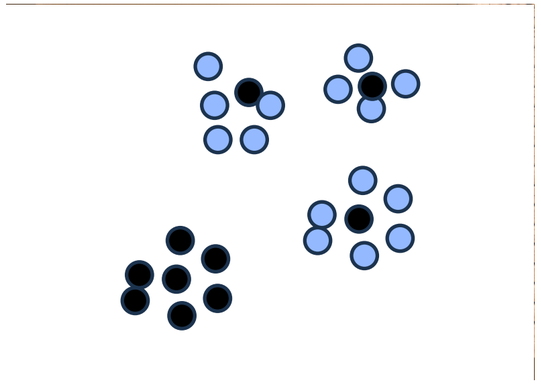


CLUSTERING

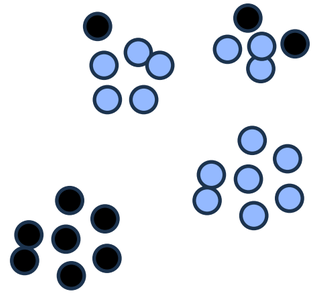
Wang et al. in preparation.

chemical space

EXPLO-RATION



CLUSTERING

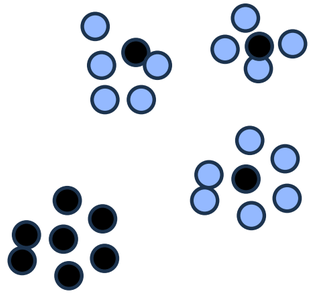


UNCERTAINTY

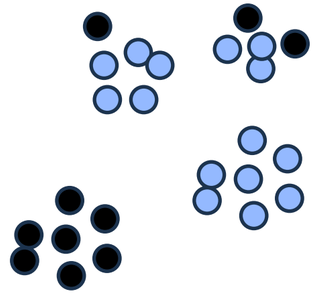
Wang et al. in preparation.

chemical space

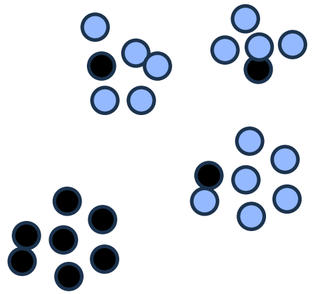
EXPLO-RATION



CLUSTERING



UNCERTAINTY

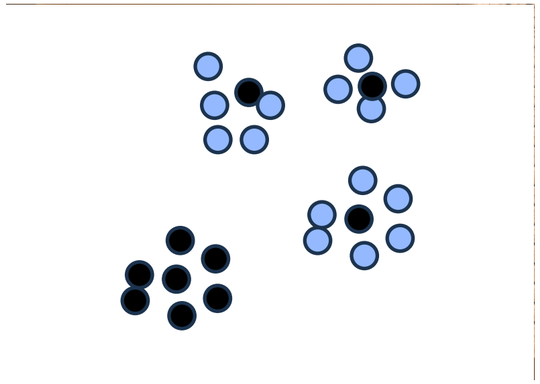


ANTICLUSTERING

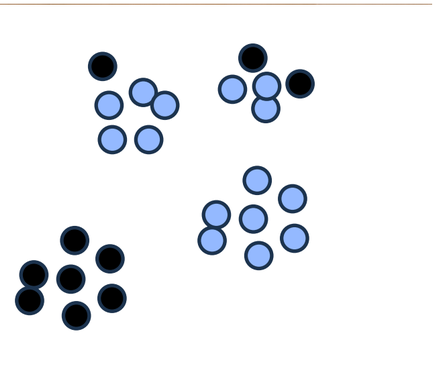
Wang et al. in preparation.

chemical space

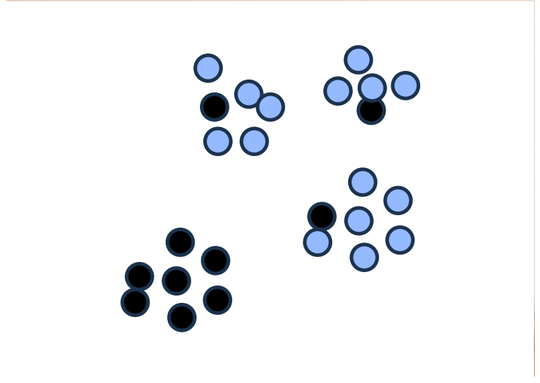
EXPLO-RATION



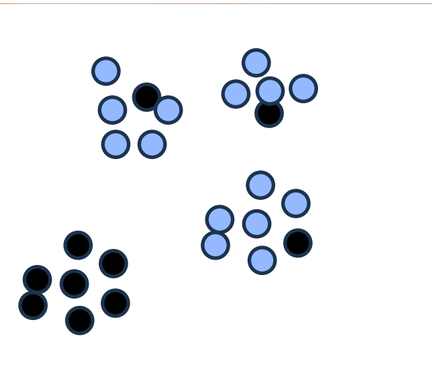
CLUSTERING



UNCERTAINTY



ANTICLUSTERING

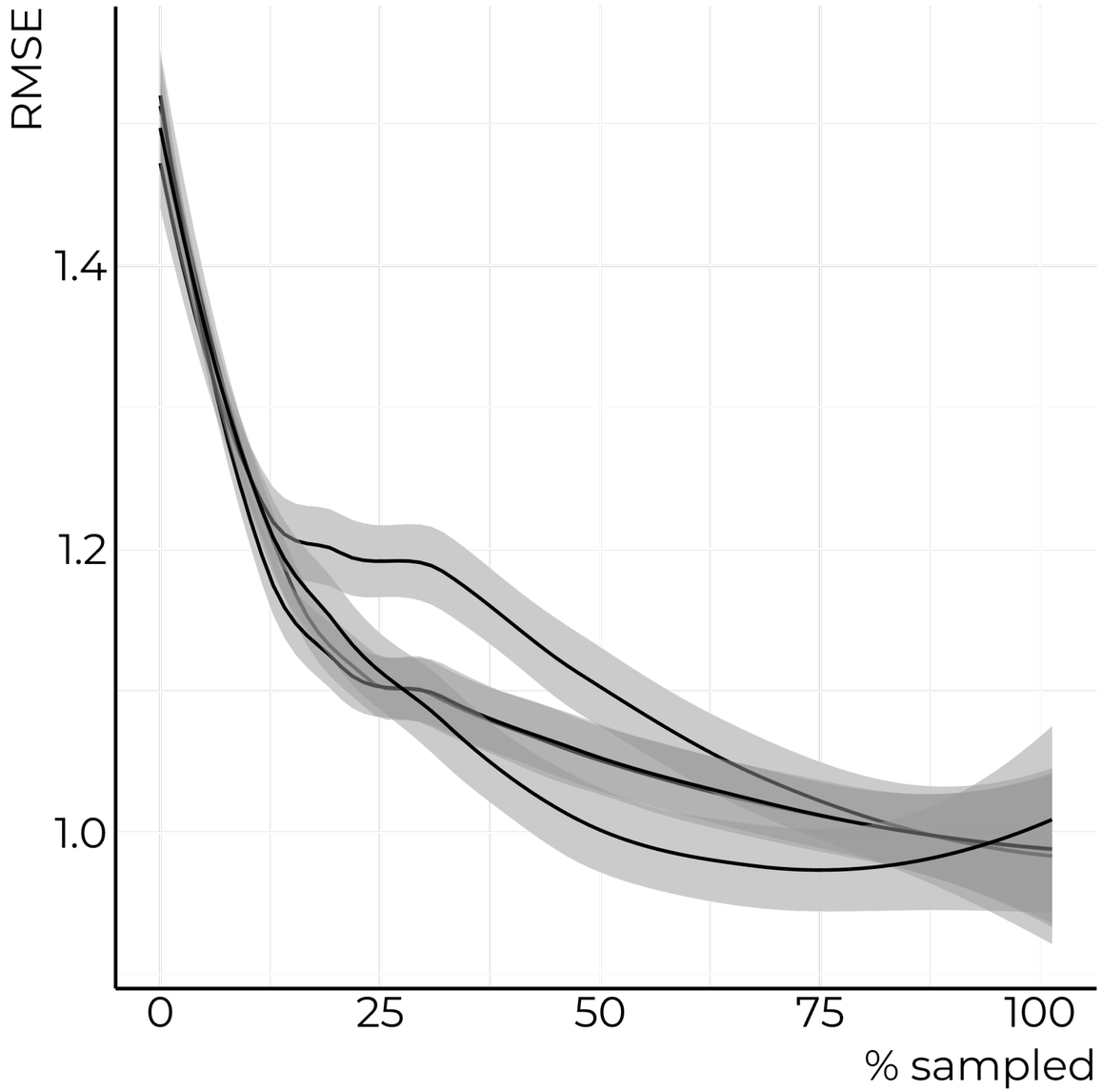


RANDOM

Wang et al. in preparation.

chemical space

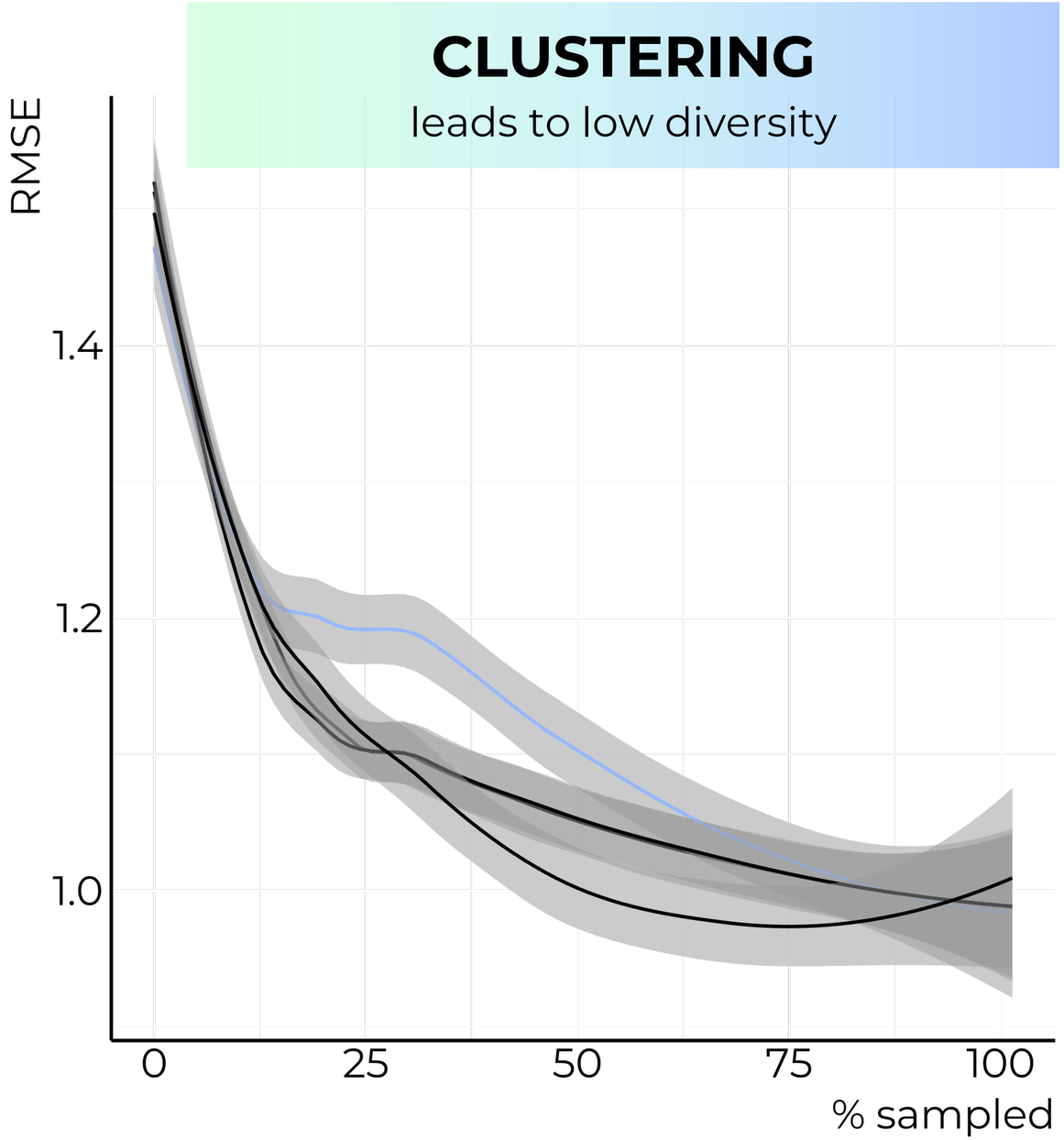
EXPLO-RATION



Wang et al. in preparation.

chemical space

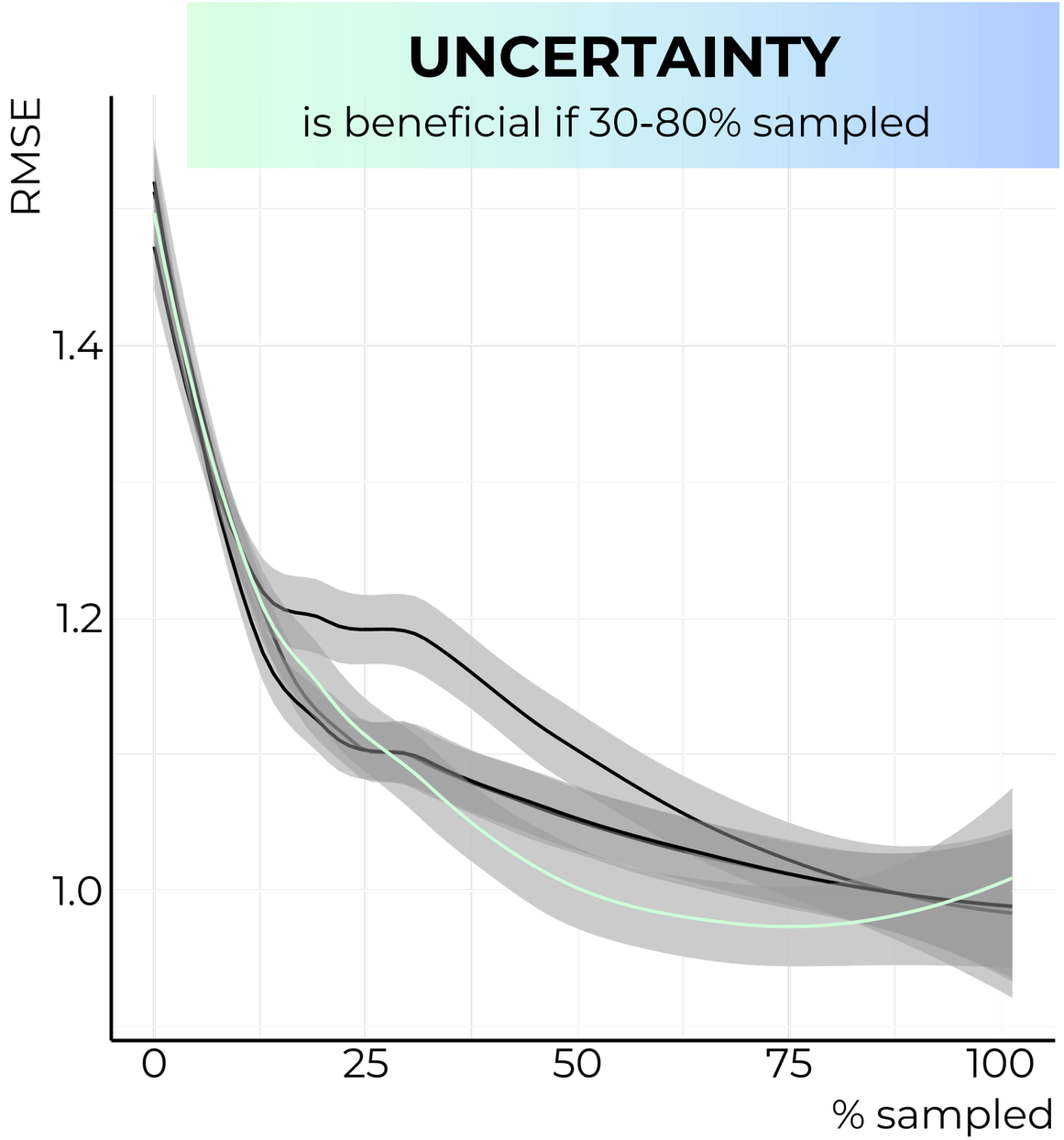
EXPLO-RATION



Wang et al. in preparation.

chemical space

EXPLO-RATION



Wang et al. in preparation.

open

QUESTIONS

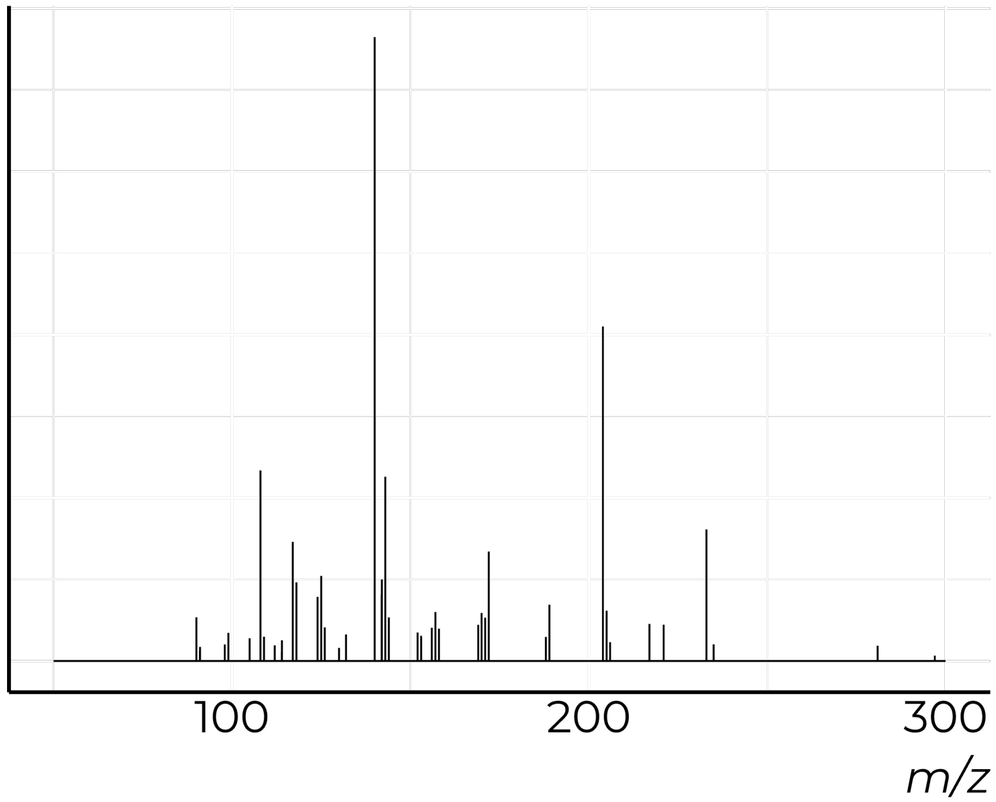
Where should we go in next 3 years?

? CHEMICAL SPACE

? UNIDENTIFIED CHEMICALS

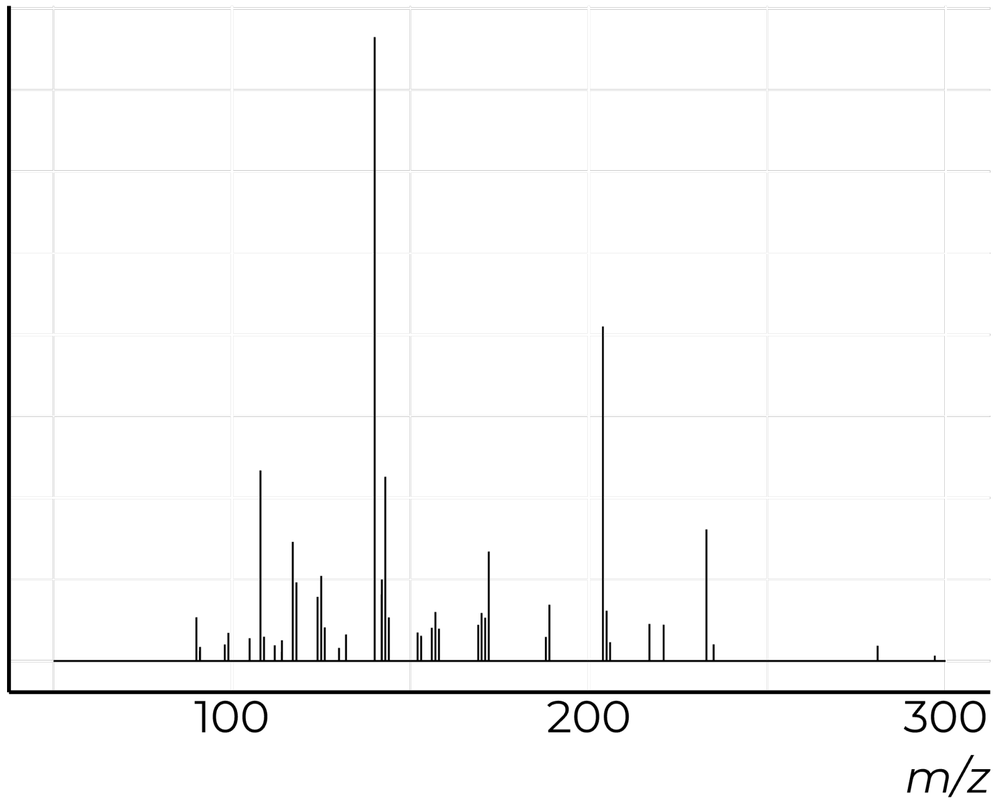
? RISK

PROBLEM

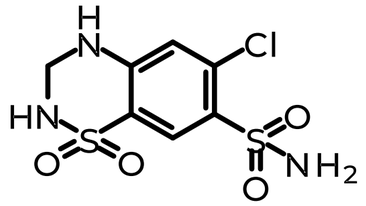


Sepman et al. Anal. Chem. **2023**, *95* (33), 12329-12338.

PROBLEM

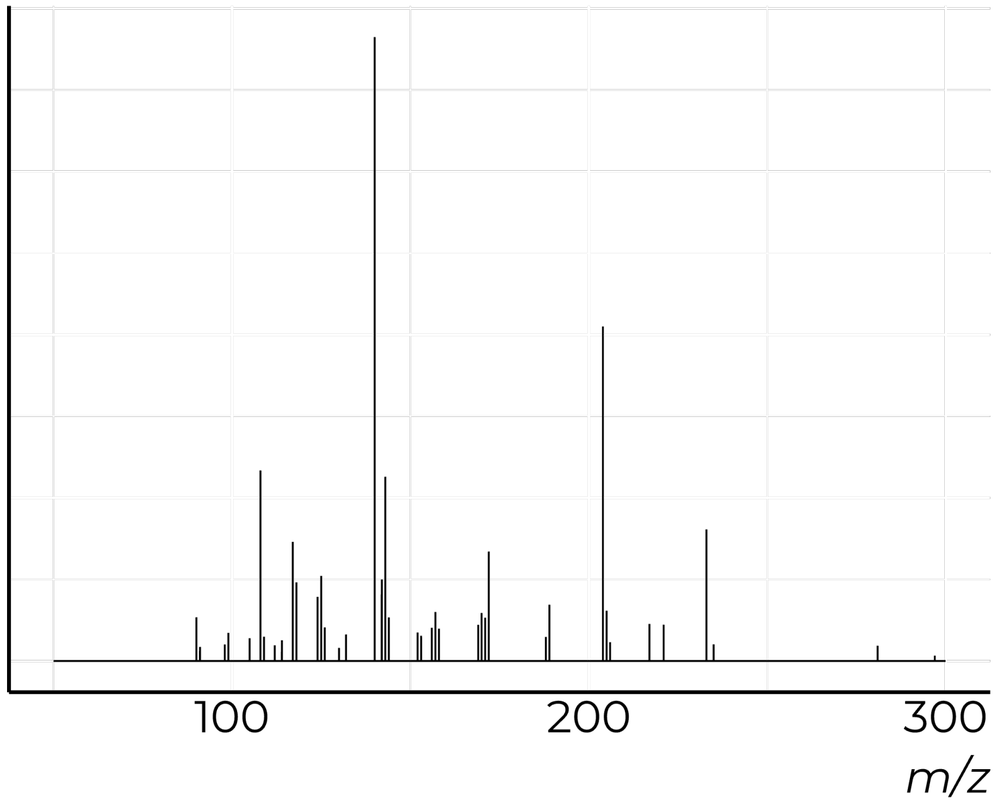


identification

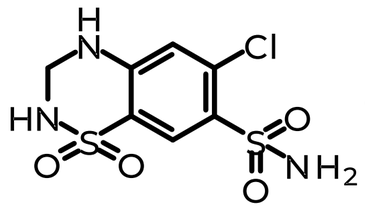


Sepman et al. Anal. Chem. **2023**, *95* (33), 12329-12338.

PROBLEM



identification →

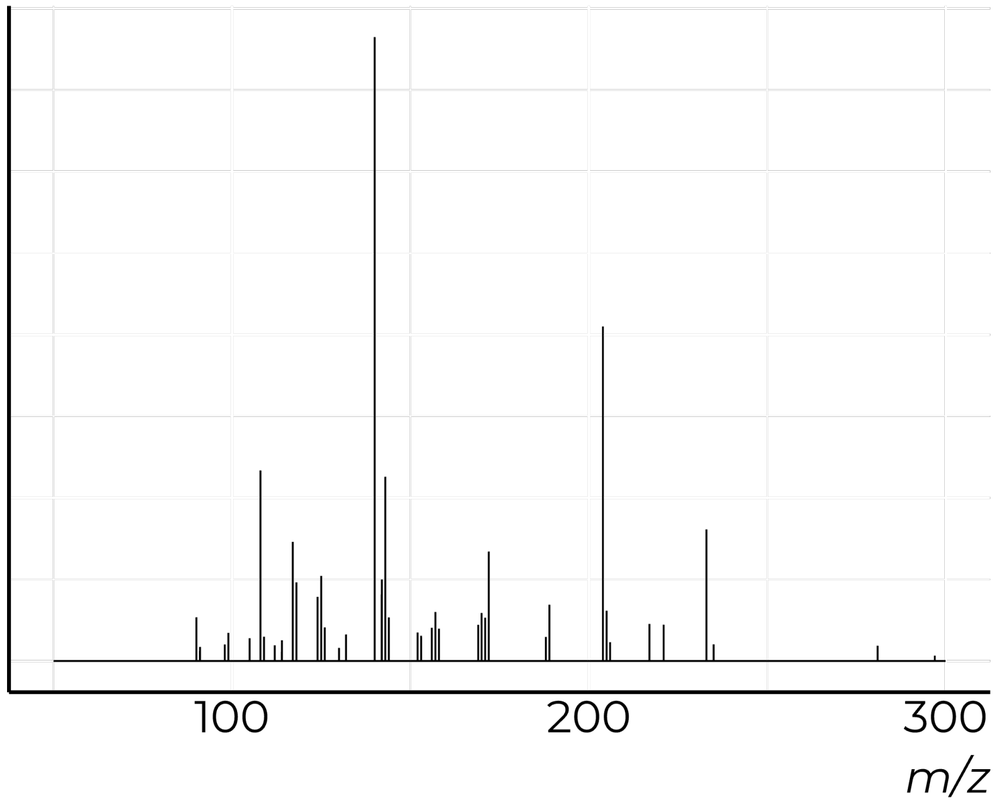


ML for IE →

c = 123 ng/L

Sepman et al. Anal. Chem. **2023**, *95* (33), 12329-12338.

PROBLEM



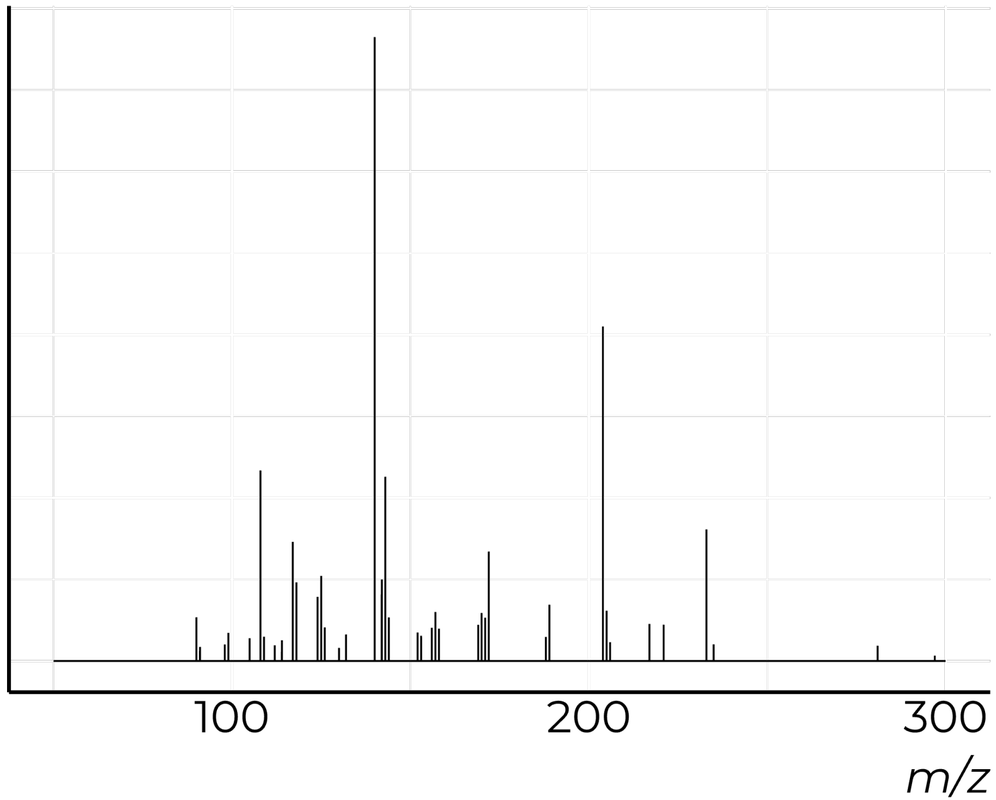
identification →

ML for IE →

c = 123 ng/L

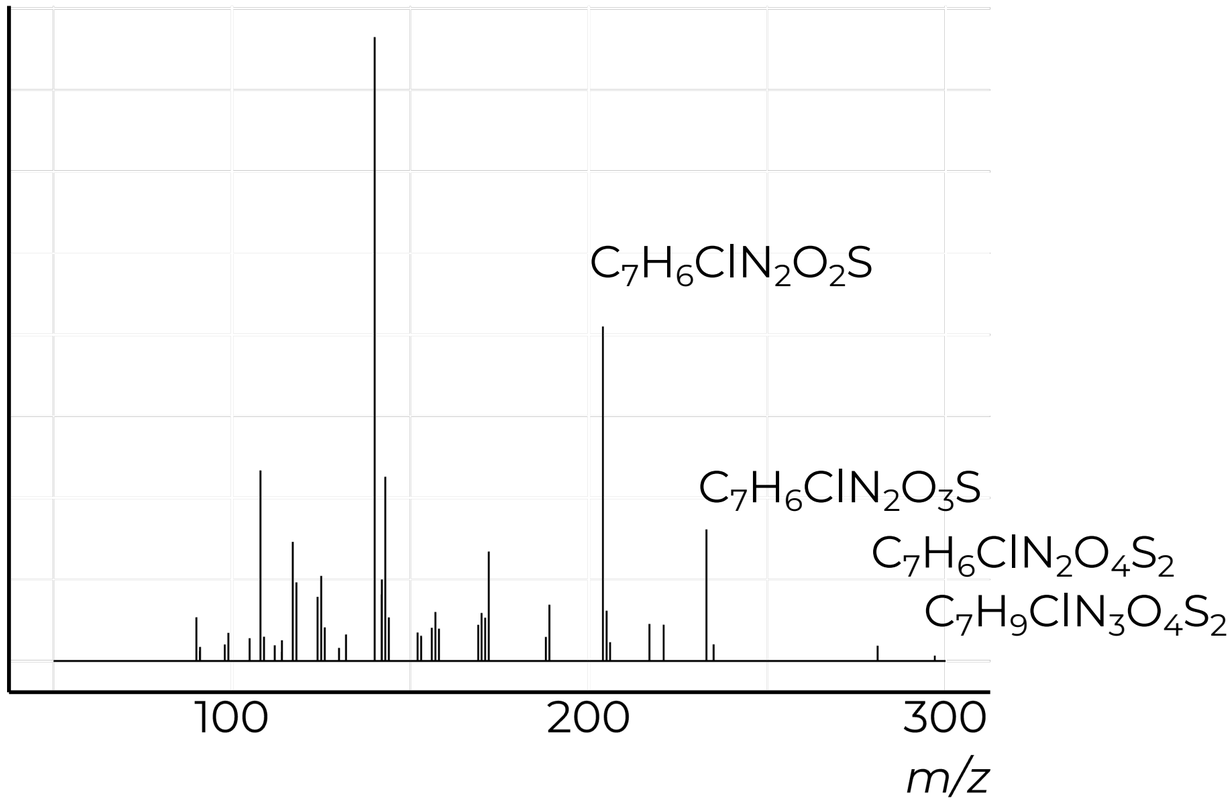
Sepman et al. Anal. Chem. **2023**, *95* (33), 12329-12338.

SOLUTION



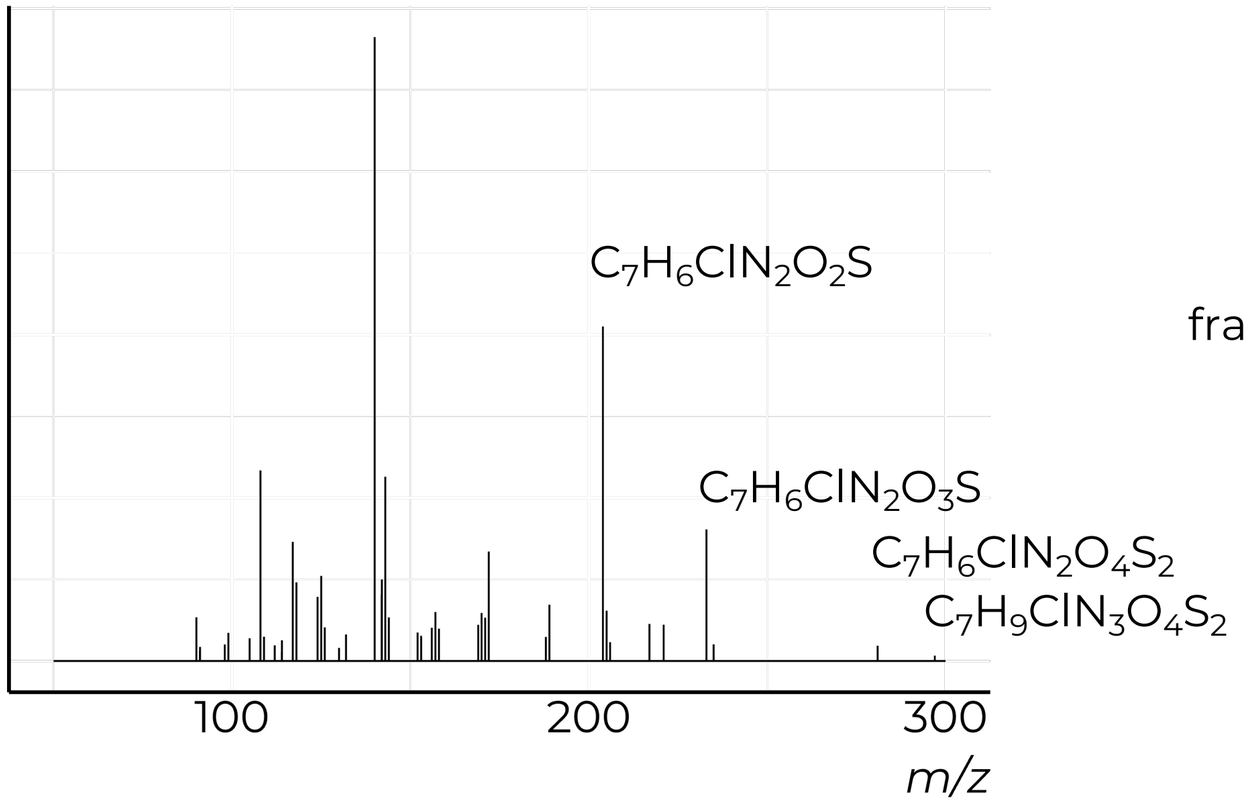
Sepman et al. Anal. Chem. **2023**, *95* (33), 12329-12338.

SOLUTION

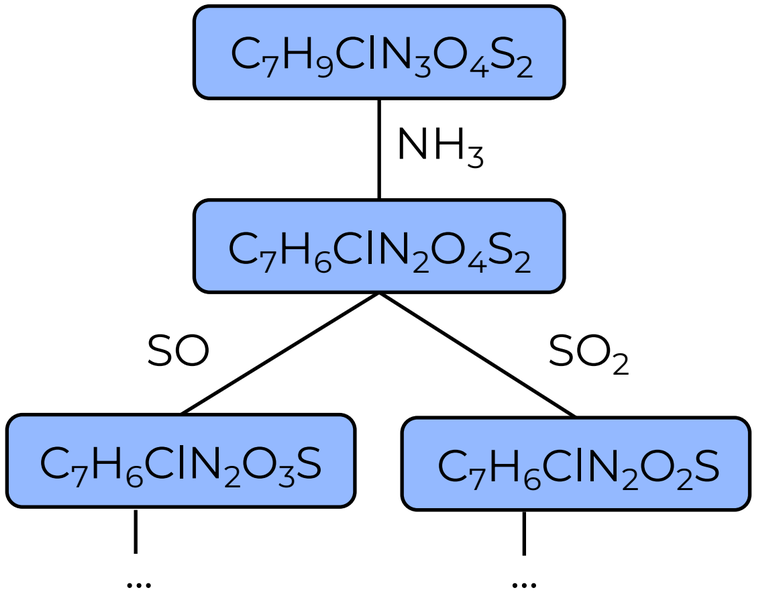


Dührkop et al. Nat. Methods **2019**, *16*, 299-302.

SOLUTION

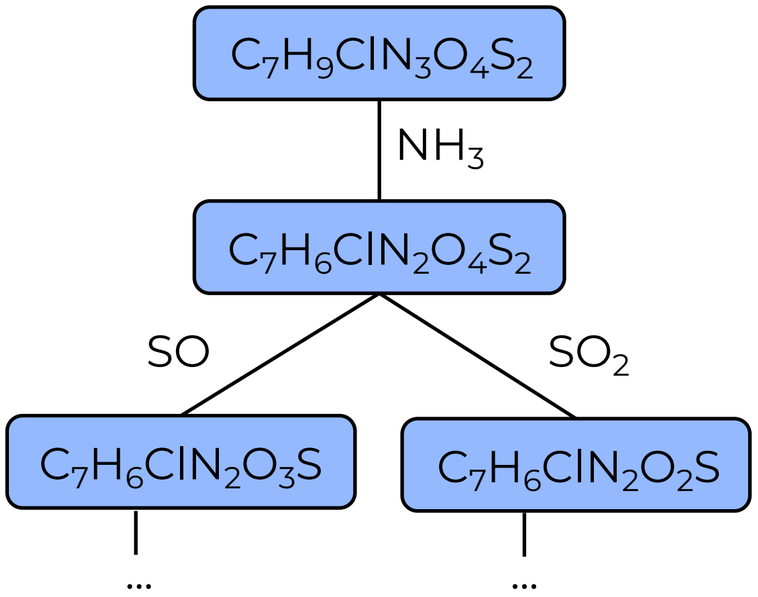


fragmentation tree



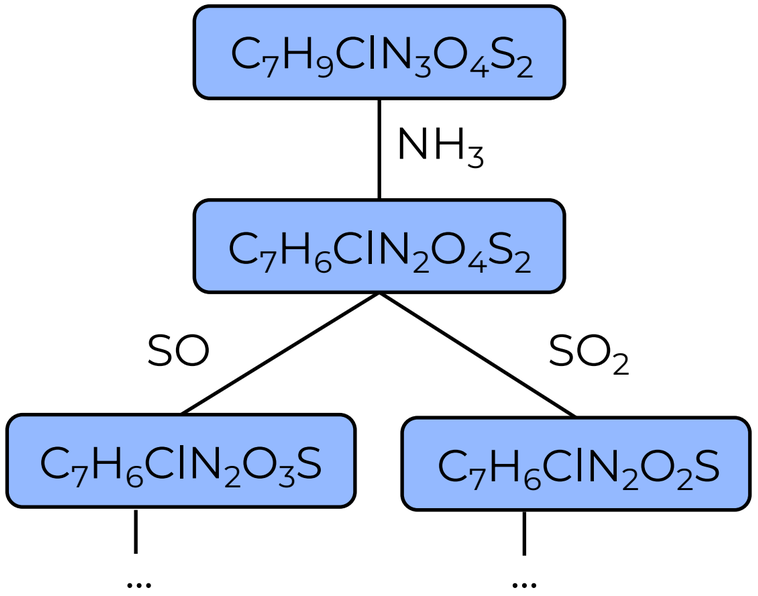
Dührkop et al. Nat. Methods **2019**, *16*, 299-302.

SOLUTION



Dührkop et al. Nat. Methods **2019**, *16*, 299-302.

SOLUTION

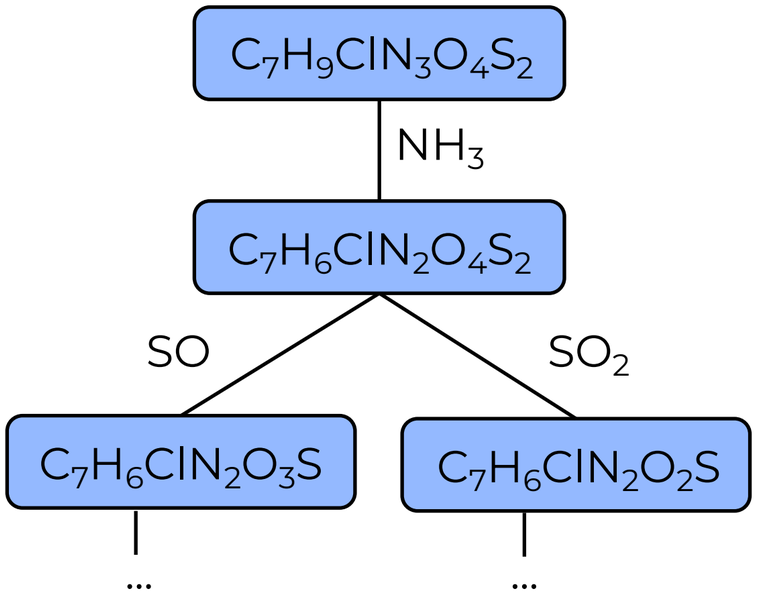


SIRIUS
CSI:FingerID

0	O-P
1	S=O
0	C-F
1	NH_2
1	O

Dührkop et al. Nat. Methods **2019**, *16*, 299-302.

SOLUTION



SIRIUS CSI:FingerID →

0	O-P
1	S=O
0	C-F
1	NH_2
1	O

ML for IE & recalibration →

c = 123 ng/L

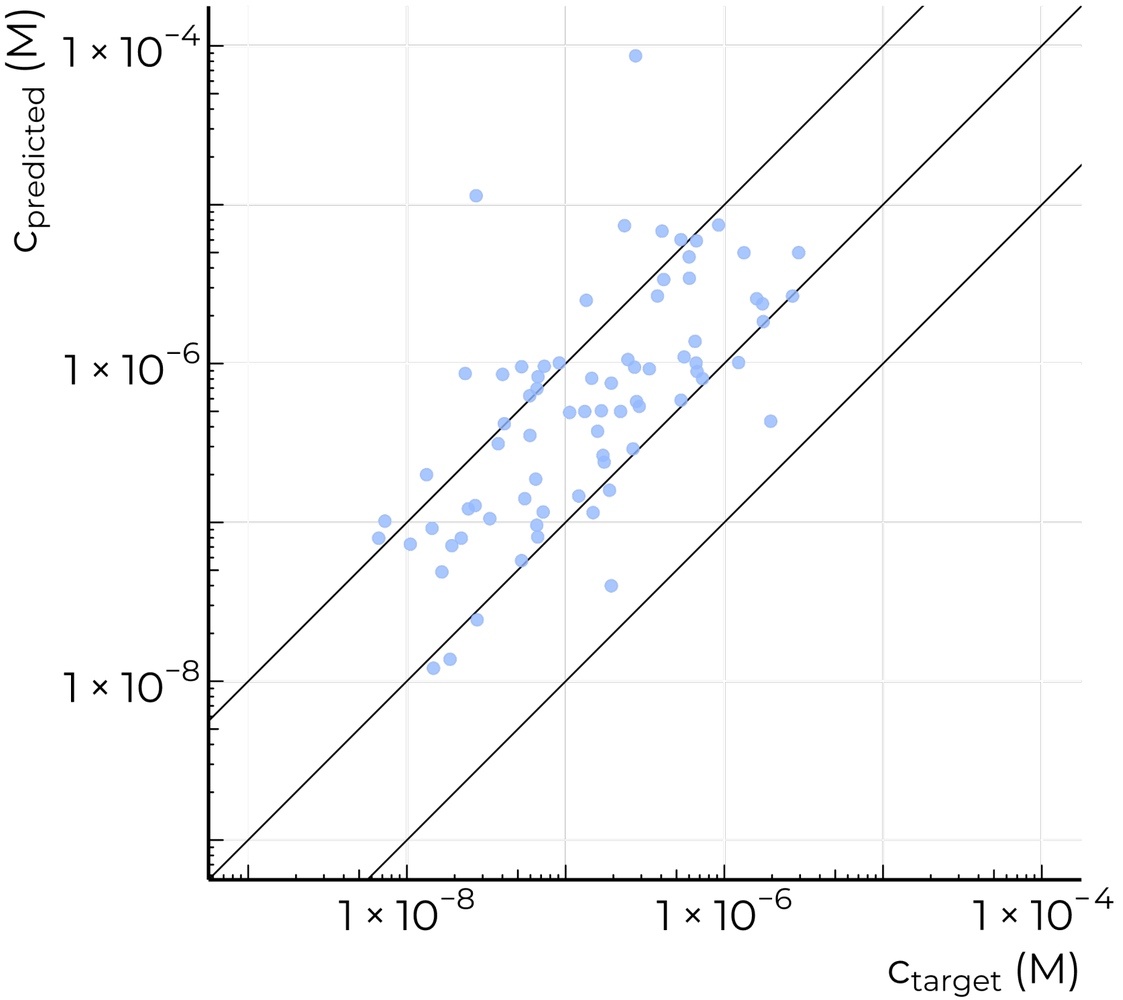
Sepman et al. Anal. Chem. **2023**, *95* (33), 12329-12338.

UNIDENTIFIED CHEMICALS

39 environmental pollutants

spiked surface water samples

median error 4.0x



Sepman et al. Anal. Chem. **2023**, *95* (33), 12329-12338.

open

QUESTIONS

Where should we go in next 3 years?

? CHEMICAL SPACE

? UNIDENTIFIED CHEMICALS

? RISK

it all boils down to

RISK

Highly toxic and high exposure substances deserve attention

TOXICITY

CONCENTRATION

it all boils down to

RISK

Highly toxic and high exposure substances deserve attention

TOXICITY

CONCENTRATION

RISK

$$risk = \frac{concentration}{toxicity}$$

toxicity of

UNKNOWNS

From MS^2 spectra

FISH LC_{50}

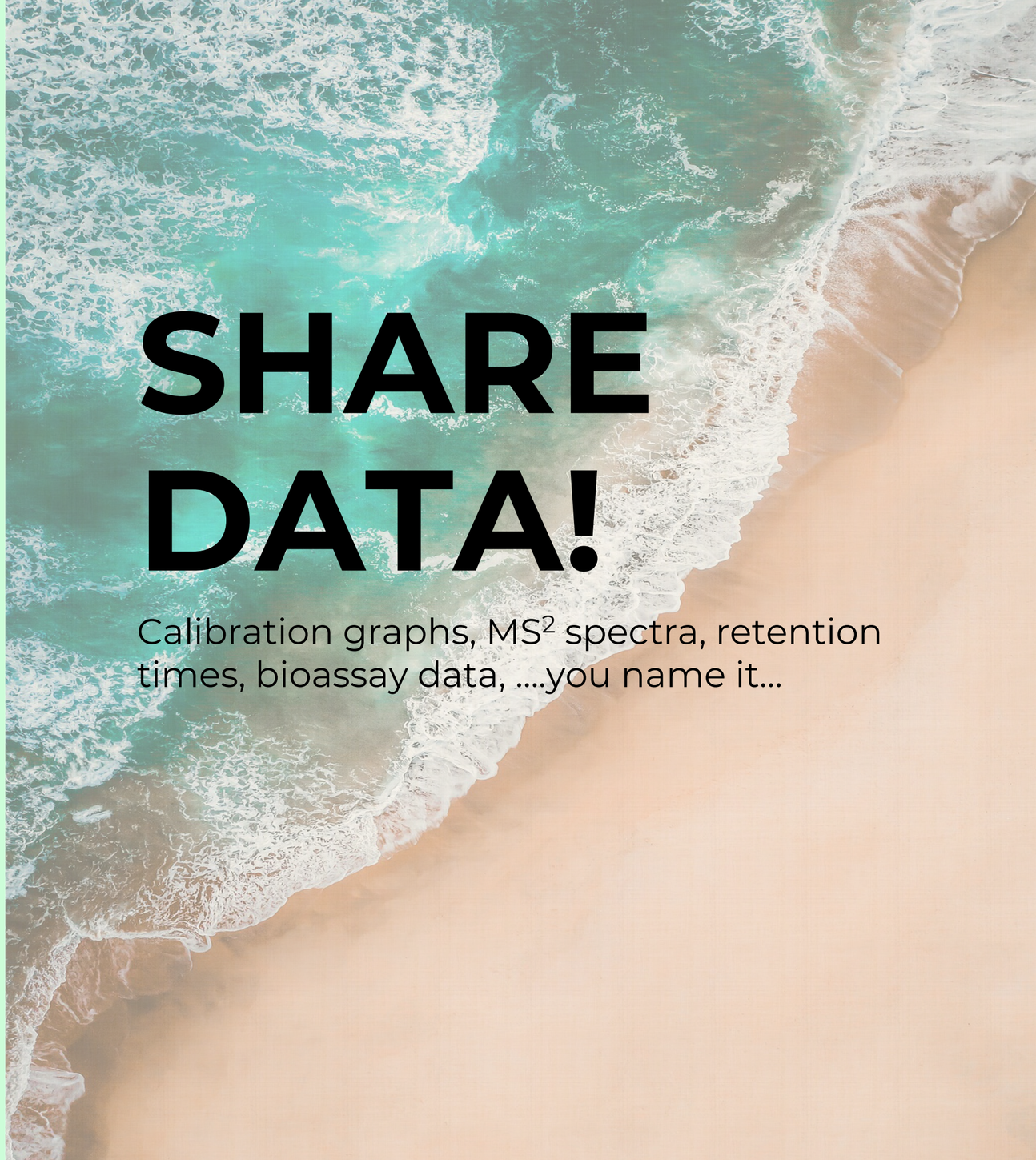
Peets et al. ES&T **2022**, *56*, 15508−15517.

ENDOCRINE DISRUPTORS

Rahu et al. JCIM **2024**, *64* (8), 3093-3104.

way forward

IN NTS?

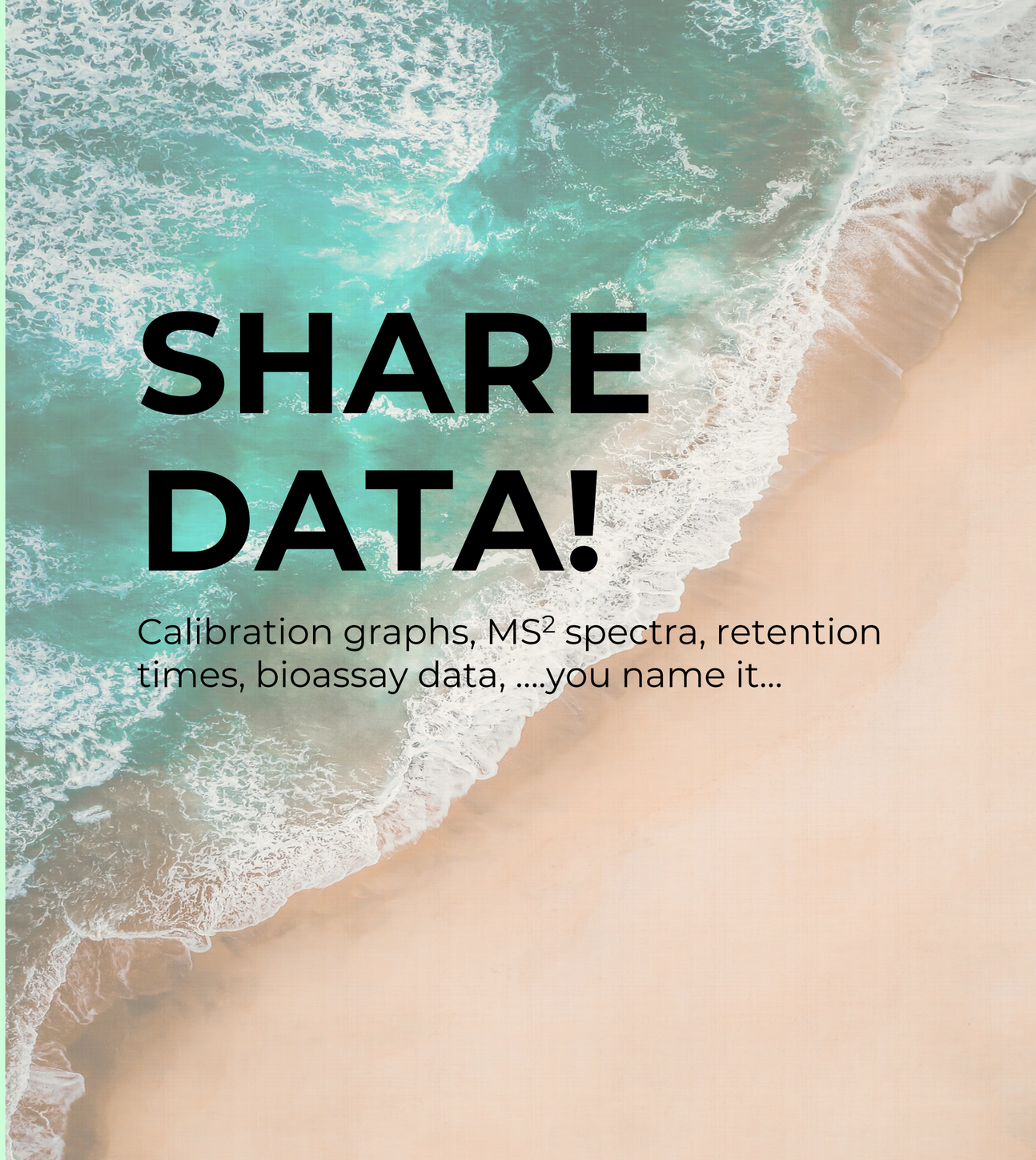


SHARE DATA!

Calibration graphs, MS^2 spectra, retention times, bioassay data, ...you name it...

SETAC Europe

3.09 – Future of Suspect and Non-Target Screening to Monitor Emerging Contaminants in the Environment



SHARE DATA!

Calibration graphs, MS^2 spectra, retention times, bioassay data, ….you name it…

anneli kruve
anneli.kruve@su.se