SETAC Europe 35th Annual Meeting – 2025 – Session 3.07.C.
Wide-Scope Target and Non-Target Screening Strategies for Enhanced Chemical Coverage in Environmental Monitoring and Chemical Exposome Assessment

Quantification without reference standard: How semi-quantitative tools performed in water sample?

Eloi Marilleau[1,3], Yvan Gru[2], Ronan Colin[2], Nicolas Cimetiere[3]

[1] Inovalys, F-56000, Vannes, France

[2] Inovalys, F-44300, Nantes, France

[3] Univ Rennes, ENSCR, CNRS, ISCR- UMR 6226, F-35000, Rennes, France

Introduction: Case study

French drinking water:

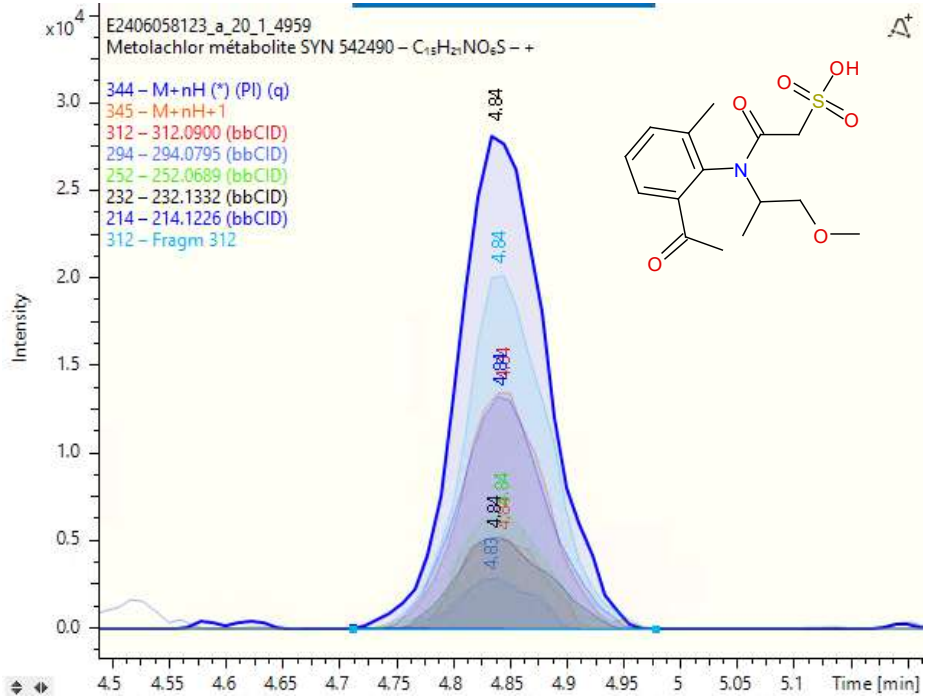
- Metolachlor (<< 0.01 µg/L)
- Metolachlor ESA (Outside calibration range > 2 µg/L)
- Metolachlor NOA 413173 (0.14 µg/L)
- Metolachlor OA (0.03 µg/L)
- **Metolachlor SYN 542490 (??? µg/L)**

Identification of metolachlor SYN 542490:

- Consistent retention time
- Low mass deviation (< 5 ppm)
- Characteristic fragment ions (MassBank Europe)

Identification level: 2a[1]

LC-HRMS chromatogram of metolachlor SYN 542490



- **Not currently included in drinking water quality** monitoring in France
- Already observed in **groundwater** (identification level 3)[2]
- Classified as **relevant** by the European Food Safety Authority (EFSA) for human health[3]

[1] E. Schymanski et al. Environ. Sci. Technol, 48 (2014) 2097–2098
[2] K. Kiefer et al. Water Research 196 (2021) 116994
[3] F. Alvarez et al. EFSA Journal 21(2):7852 (2023) 46p

Introduction: Case study

French drinking water:

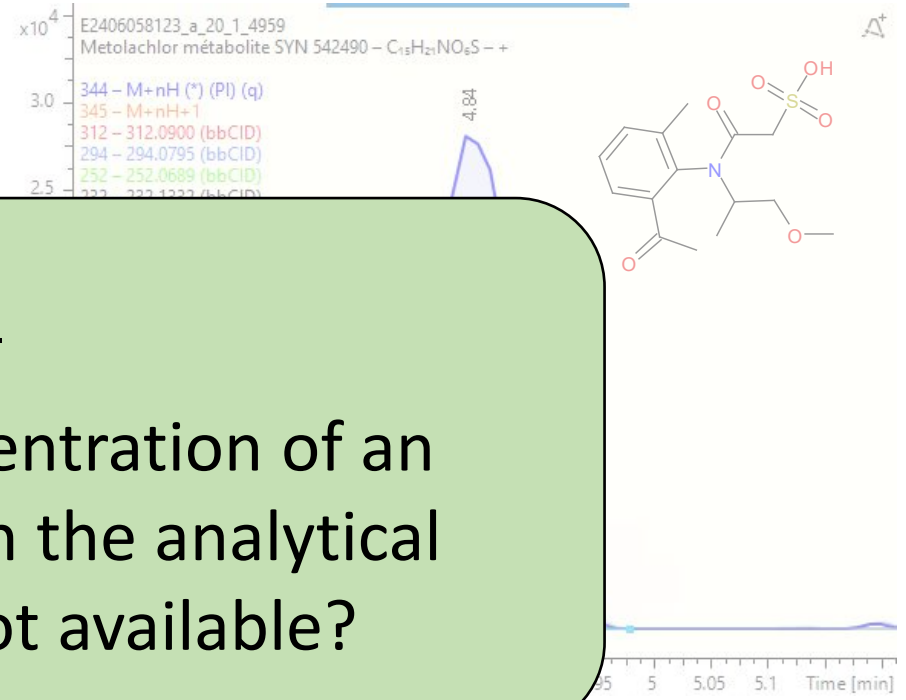
- Metolachlor (<< 0.01 µg/L)
- Metolachlor ESA (Outside calibration range > 2 µg/L)
- Metolachlor NOA 413173 (0.14 µg/L)
- Metolachlor O
- Metolachlo

Identification of me

- Consistent r
- Low mass d
- Characterist

Identification level: 2a[1]

LC-HRMS chromatogram of metolachlor SYN 542490



Objectives :

How to estimate the concentration of an identified compound when the analytical reference standard is not available?

- **Not currently included in drinking water quality** monitoring in France
- Already observed in **groundwater** (identification level 3)[2]
- Classified as **relevant** by the European Food Safety Authority (EFSA) for human health[3]

[1] *E. Schymanski et al. Environ. Sci. Technol, 48 (2014) 2097–2098*
[2] *K. Kiefer et al. Water Research 196 (2021) 116994*
[3] *F. Alvarez et al. EFSA Journal 21(2):7852 (2023) 46p*

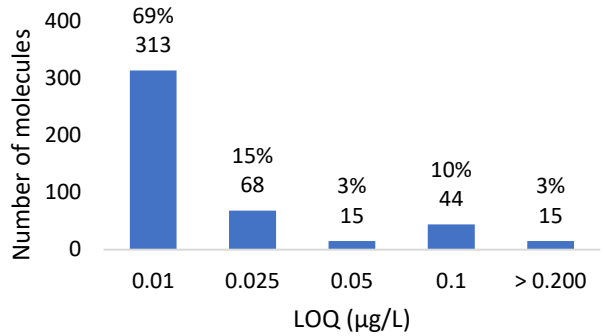
Workflow: Water analysis

0.7 µm glass fiber filtration
+ internal standards spiking

↓

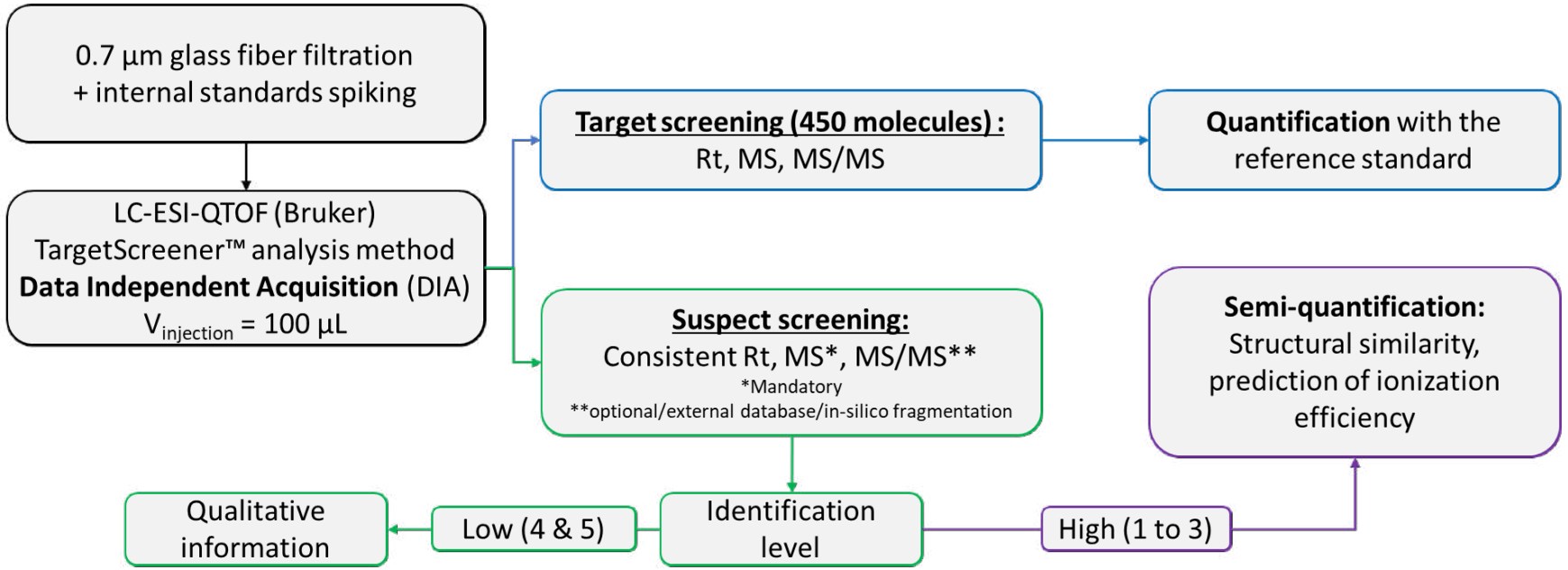
LC-ESI-QTOF (Bruker)
TargetScreener™ analysis method
Data Independent Acquisition (DIA)
$V_{injection}$ = 100 µL

Figure 1: *LC-HRMS Limits of quantification (LoQ) for 455 compounds in drinking water (ESI + et ESI -)*



E. Schymanski et al. Environ. Sci. Technol, 48 (2014) 2097−2098

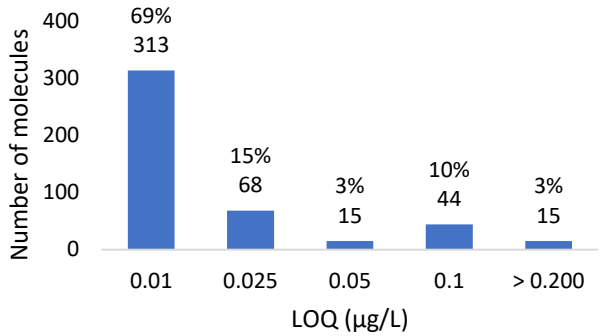
Workflow: Water analysis



4	Molecular formula	MS isotope/adduct
5	Exact mass	Δm/z < 5 ppm

1	Confirmed structure	Reference standard
2.a	Probable structure	MS^2 librairies
2.b		MS^2 exp. consistent
3	Tentative candidates	Theorical MS^2, Consistent Rt

Figure 1: *LC-HRMS Limits of quantification (LoQ) for 455 compounds in drinking water (ESI + et ESI -)*



E. Schymanski et al. Environ. Sci. Technol, 48 (2014) 2097−2098

Semi-quantification: General information

Semi-quantitative approaches aim to **estimate the concentration** of a compound based on its properties and signal intensity, **without using the reference standard**.

Retention time

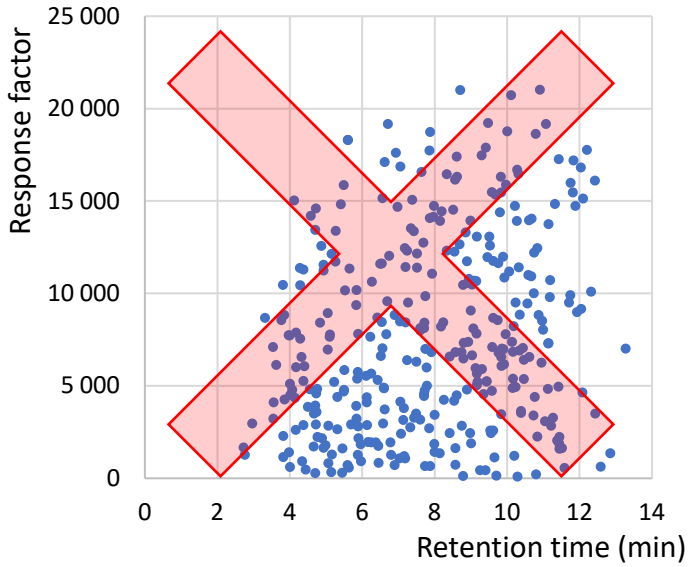
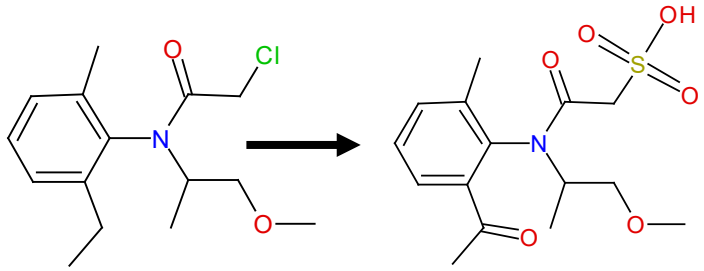


Figure 2: *Response factor versus retention time for 351 compounds*

Structural and chemical similarities

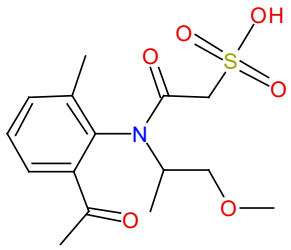
Metolachlor

Metolachlor SYN 542490



$$C_{SYN542490} = \frac{Area_{SYN542490}}{Slope_{Metolachlor}} = ?\,\mu g/L$$

Prediction of ionization efficiency



Metolachlor SYN 542490	
MW (g/mol)	343.4
-H donnor	1
-H Acceptor	6
XLogP3	0.5
…	…

Predicted response factor

$$RFpred_{SYN542490}$$

$$C_{SYN\ 542490} = \frac{Area_{SYN542490}}{RFpred_{SYN542490}}$$

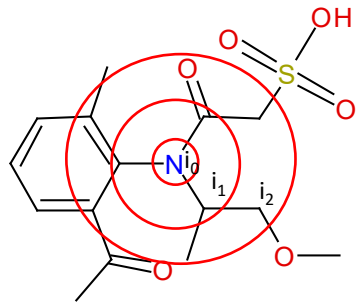
Semi-quantification: Structural and chemical similarities

Fingerprints:

- Output data: bit string encoding structural information of the molecule

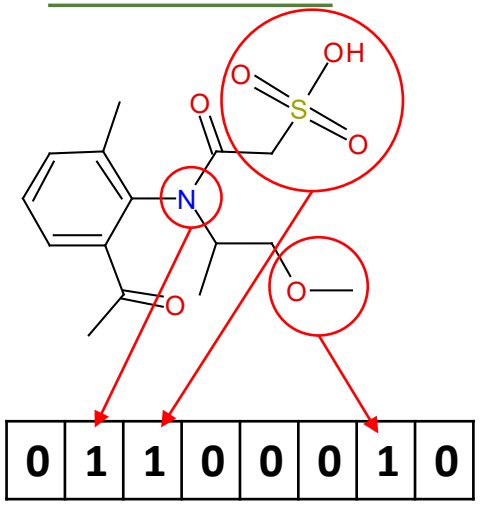
Type	Fingerprint 1	Fingerprint 2	Iterations
Circular	ECFP	FCFP	1 or 2
Substructure	MACCS	Pubchem	-

Circular



***ECFP**: Extended connectivity*
***FCFP**: Functionnal class*

Substructure



Comparison database:

- 350 compounds (pesticides, drugs, metabolites)
- Spiked in drinking water

Similarity calculation:

- Tanimoto index:

$$S_{A,B} = \frac{c}{a + b - c}$$

$S_{A,B}$: Similarity [0;1] between molecule A and B
a : Activated bit in A
b : Activated bit in B
c : Activated bit in A and B

Concentration calculation:

$$C_{Suspect} = \frac{Area_{suspect}}{slope_{\ \mathbf{Top\ 1}\ similar}}$$

Semi-quantification: Prediction of ionization efficiency

How it works?

- **Machine learning tools**
- Use physico-chemical descriptors of molecules (logP, number of atoms, -H bond acceptor, -H bond donor, ...)
- Prediction of the concentration

Quantem

Model : Random forest

ESI positive:

- $[M+H]^+$, $[M]^+$
- **106** mobile phase compositions
- **353** compounds
- **450** descriptors

ESI negative:

- $[M-H]^-$
- **33** mobile phase compositions
- **101** compounds
- **145** descriptors

Quantem Analytics: https://quantem.co/
J. Liigand et al. Scientifique reports (2020) 10:5808

Semi-Quantification of Emerging Pollutants (SQEP, v. 3.0)

Model: Ant colony optimization – multilinear regression

ESI positive:

- $[M+H]^+$, $[M]^+$
- Mobile phase composition not considered
- **106** compounds
- **7** descriptors

SQEP: https://trams.chem.uoa.gr/semiquantification-2/
R. Aalizadeh et al. Anal Bioanal Chem 414 (2022) 7435–7450

Semi-quantification: Comparison and validation methodology

Comparison

Molecule selection criteria:

- ESI+ ionizable
- **[M+H]⁺** or **[M]⁺** only

Summation of in-source fragment area to the principal ion area.

→ **351 molecules in ESI+** :

- Pesticides
- Pesticides metabolites
- Pharmaceuticals

Comparison matrices:

- Mineral drinking water spiked at 0.100 or 0.500 µg/L
- 2 analytical series for repeatability assessment

Comparison criteria:

- Error factor mean and median
- Outliers error factor (EF > 10)

$$Error\ Factor\ (EF) = MAX \begin{cases} \dfrac{Known\ concentration}{Estimated\ concentration} \\ \dfrac{Estimated\ concentration}{Known\ concentration} \end{cases}$$

Validation

Validation study:

- Performances on proficiency tests samples:
  - Surface water (n = 5)
  - Drinking water (n = 2)

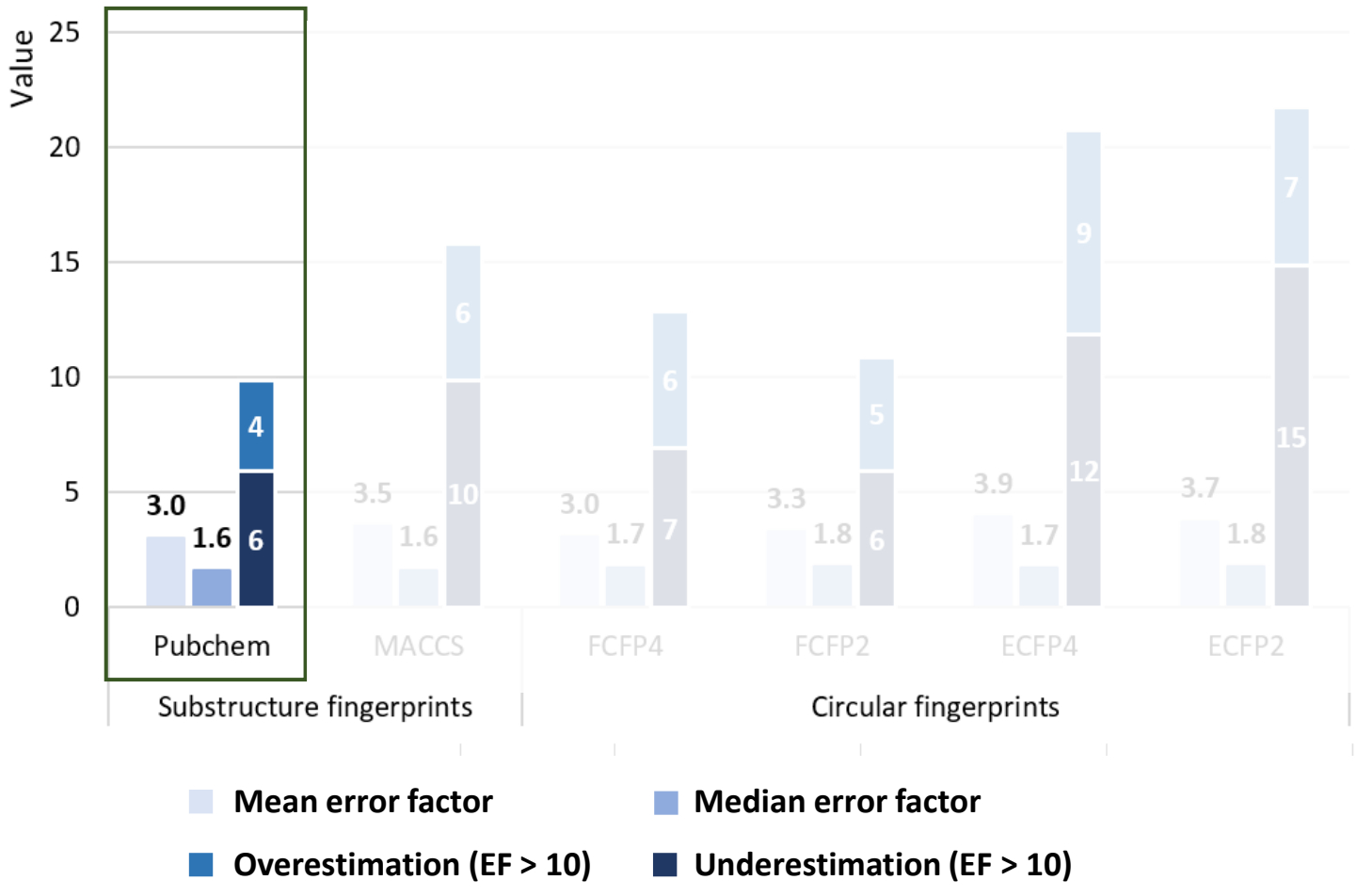
→ Pesticides and pesticides metabolites

Results and discussions

- Comparisons of semi-quantitative approaches
- Validation study
- Application to the case study

Semi-quantification: Similarity approach

Figure 3: *Mean EF, median EF and number of outliers for different chemical fingerprints for the search of TOP 1 similar compound for semi-quantification (351 molecules)*



Comparable **mean** error factors

Comparable **median** error factors

Low inter-series **variability**

Differences in **outliers** (EF > 10)

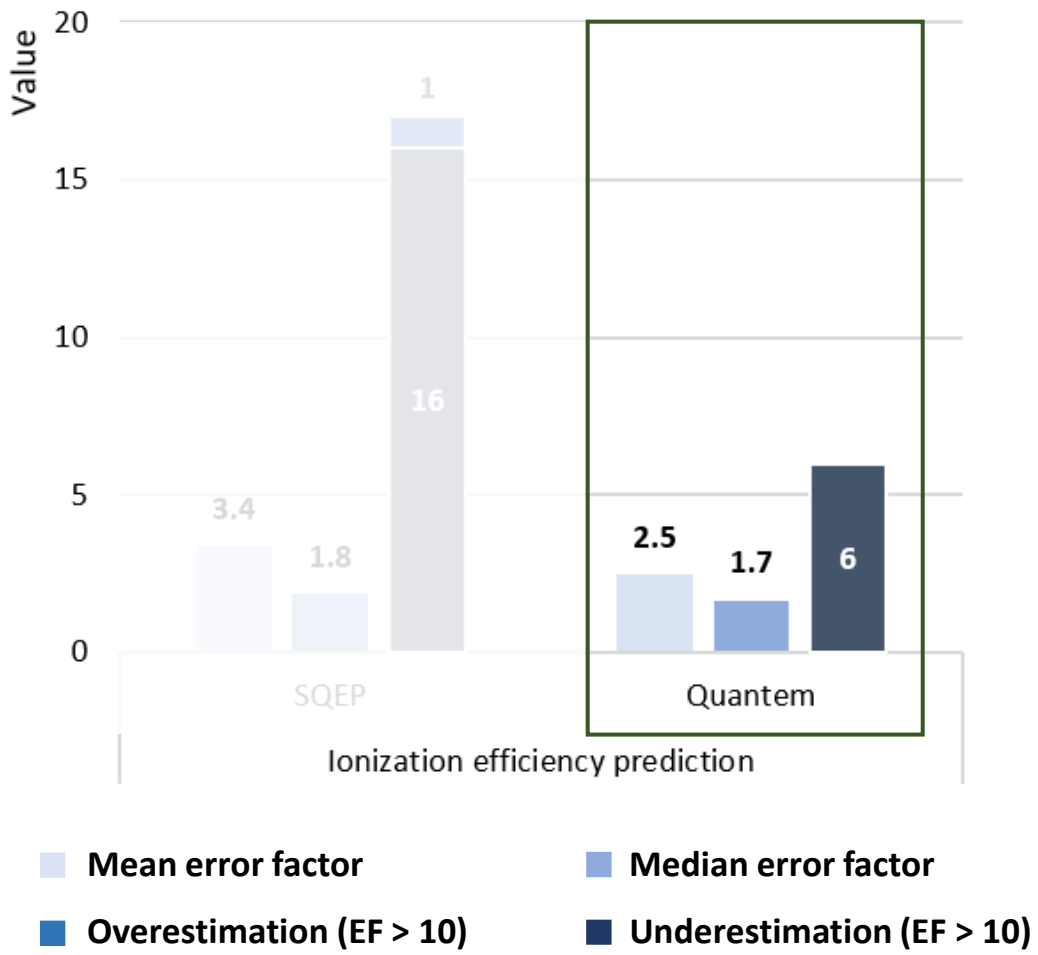
Slight **underestimation** trend

Selected fingerprint:

→ Pubchem

Semi-quantification: Prediction of ionization efficiency

Figure 4: *Mean EF, median EF and number of outliers for ionization efficiency prediction softwares (351 molecules)*



Lower **mean** error factors for Quantem

Comparable **median** error factors

Low inter-series **variability**

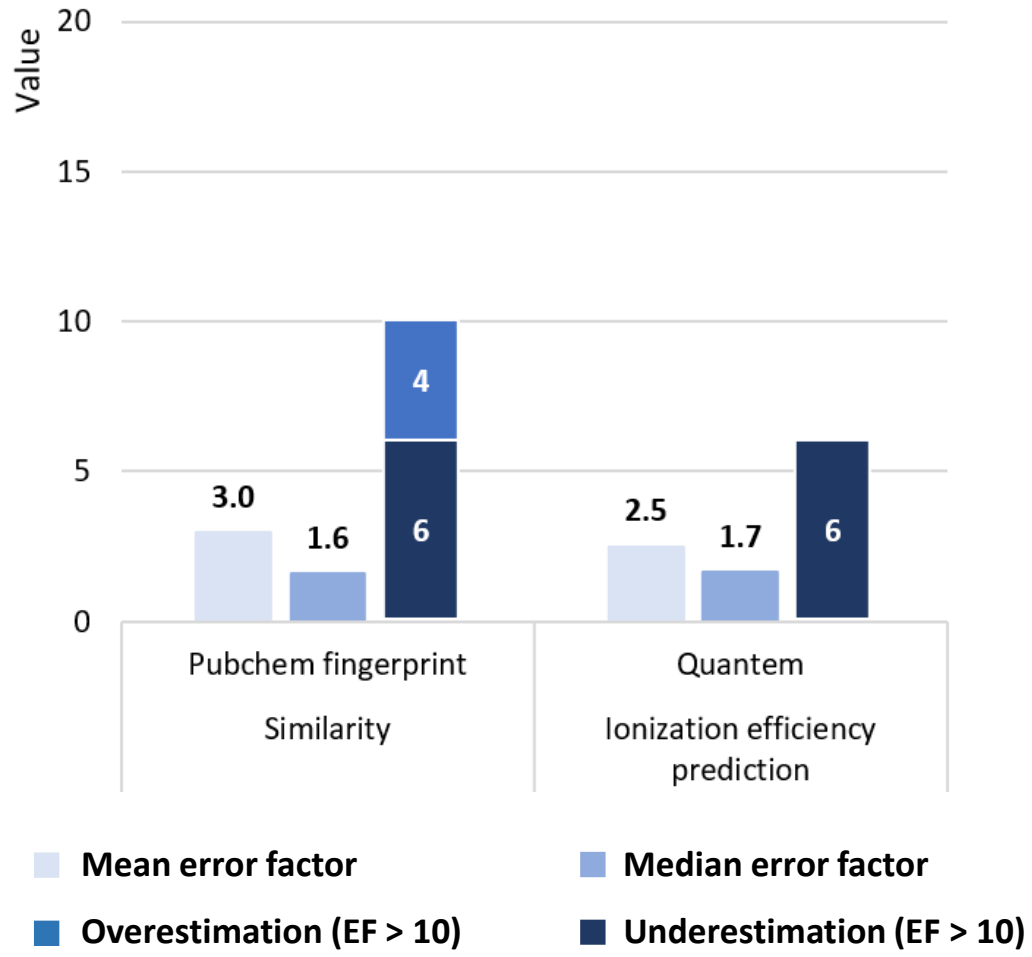
Differences in **outliers** (EF > 10)

Underestimation trend

Selected software:
→ **Quantem**

Semi-quantification: Similarity vs. Prediction of ionization efficiency

Figure 5: *Mean EF, median EF and number of outliers for Pubchem fingerprint and Quantem (351 molecules)*



Best-performing approaches

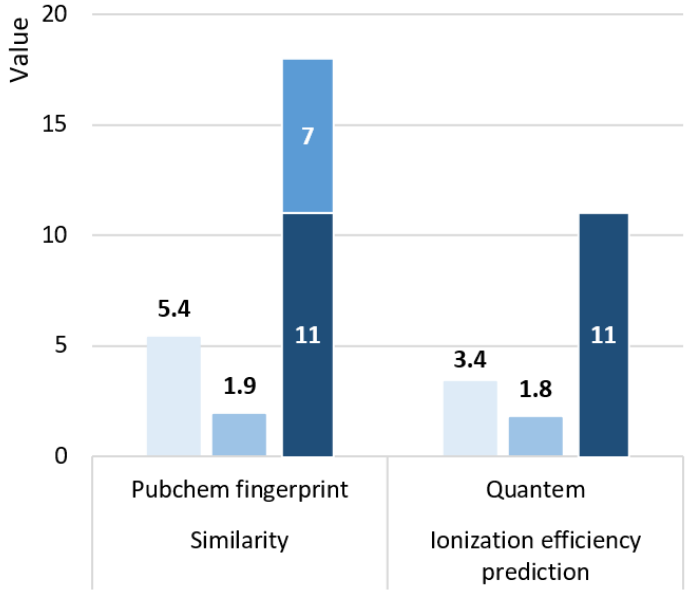
Mean error factors ≤ 3

Median error factors < 2

Lowest number of outliers

Application to proficiency tests

Figure 6: *Mean EF, median EF and outliers results for 202 molecules in proficiency test*



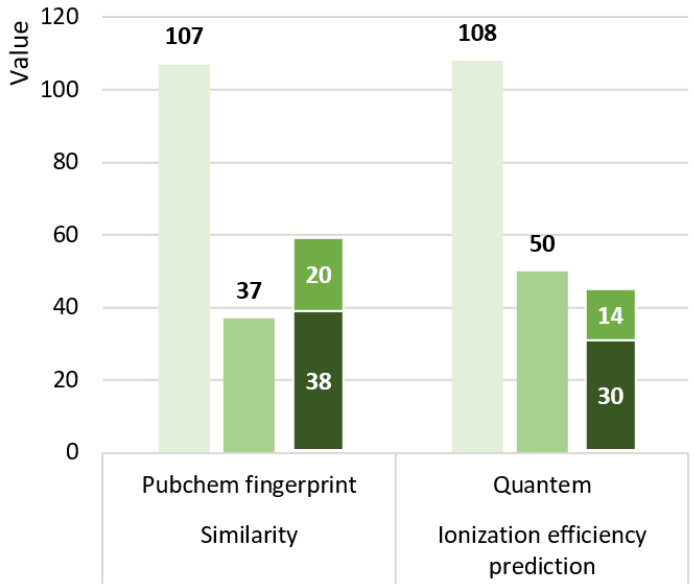
Mean error factor

Median error factor

Overestimation (EF > 10)

Underestimation (EF > 10)

Figure 7: *z-score results for 202 molecules in proficiency test*



Satisfactory (|z-score| < 2)

Questionable (2 < |z-score| < 3)

Overestimation (z-score > 3)

Underestimation (z-score < - 3)

$$z - score = \frac{C_{exp.} - C_{Assigned}}{\sigma_{prof.assess.}}$$

- C_{exp}: experimental concentration
- $C_{assigned}$: assigned value by the organism
- $\sigma_{prof.assess}$: standard deviation for proficiency assessment

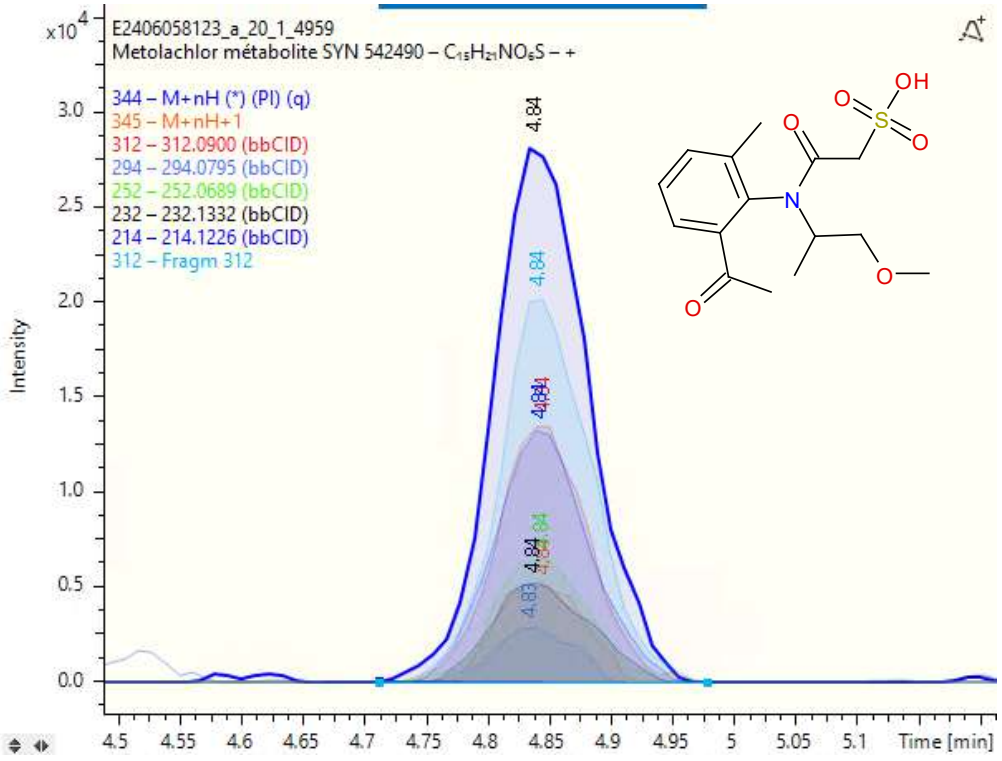
Proficiency tests information:

- Surface water (5) and drinking water (2)
- Organic contaminants (pesticides and metabolites)
- 202 semi-quantified compounds

> 70% of compounds with **acceptable z-score (|z-score| < 3)**

Case study: Metolachlor metabolite SYN 542490

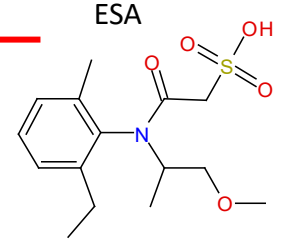
LC-HRMS Chromatogram of Metolachlor SYN 542490



Drinking water sample:

Name	Identification level	SQ Quantem (µg/L)	SQ Pubchem (µg/L)
Metolachlor SYN 542490	2a	0.15	0.35

Top 1:
Metolachlor ESA



Other detections:

French water samples (n = 97)	
Detection frequency Inlet drinking water treatment plant	Detection frequency Outlet drinking water treatment plant
30/48 (**63%**)	29/49 (**59%**)

Conclusions and perspectives

Comparison of semi-quantitatives approaches:

Prediction of ionization efficiency

- ✓ Best results with Quantem
- ✓ Fast
- ✓ Easy to set up
- ✓ Few analytical standard needed
- X Paid software (Quantem)

Similarity approach

- ✓ Best results with Pubchem fingerprints
- ✓ Fast
- ✓ No additional costs*
- X *Requires numerous analytical standards

Perspectives:

- Correction of matrix effect using isotopically labelled internal standard
- Application to other samples (wastewater, soil, food, feed)
- Tests to be carried out in negative ionization mode

Thanks for your attention

Any questions ?

Acknowledgements:
Inovalys:
- Yvan Gru
- Ronan Colin

Rennes university:
- Nicolas Cimetiere

Contact: Eloi Marilleau
eloi.marilleau@inovalys.fr