Unifying ionization efficiencies: quantitative comparison of diverse data sets and validation of prediction models

Piia Liigand[a], Jaanus Liigand[a], Karl Kaupmees[a], Anneli Kruve[a,b]
piia.liigand@ut.ee | kruvelab.com
[a]Institute of Chemistry, University of Tartu, Estonia
[b]Department of Environmental Science and Analytical Chemistry, Stockholm University, Sweden

Overview

Purpose: Develop an approach to combine data from different sources to unified dataset.

Methods: Random forest algorithm based ionization efficiency prediction model was used to calculate predicted ionization efficiencies.

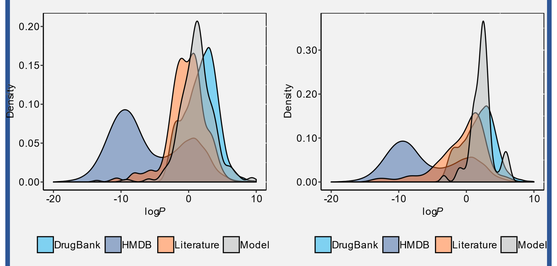
Results: All ionization efficiency data (either in the form of response factors, ionization efficiencies or signal and concentration) from past and future research can be combined into one quantitatively comparable dataset.

Motivation

Ionization efficiency. Compounds ionizing in ESI source can have **10^6** times different ionization efficiencies. Using signals to quantify with would lead to a misestimation of 1 million times.

Data is fragmented. Usually, a compound class or a few compound classes are studied. The ESI setup can vary: ESI mode, used eluents, water phase additives, etc.

Data is contradictory. Because of used limited set of variables (compounds, eluents, setup) the trends seem not to be relevant for a subset may be relevant for the whole set.



Comparison of the chemical space coverage based on log*P* values. *HMDB also includes compounds that have not been and cannot be measured with LC/MS*.

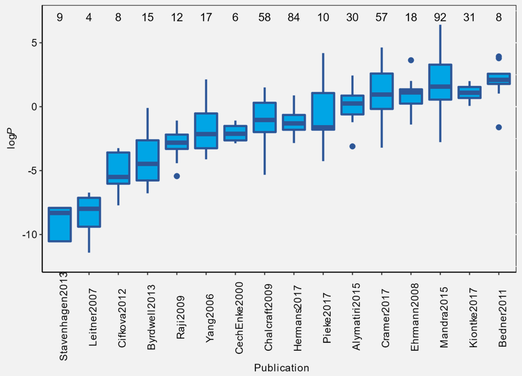
Literature sources

57 papers analyzed:
- 15 ESI positive mode,
- 7 ESI negative mode,
- 2 both modes.

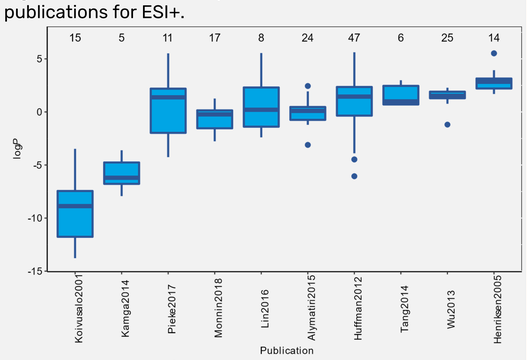
RF, RRF, IE, signal-concentration.

ESI+
634 response factors:
- 440 unique compounds,
- 64 different eluent compositions.

ESI–
373 response factors:
- 161 unique compounds,
- 47 eluent compositions.



log*P* distributions of compounds studied in individual publications for ESI+.



log*P* distributions of compounds studied in individual publications for ESI–.

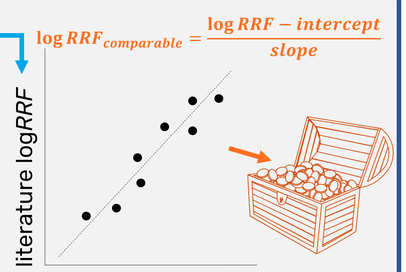
Unification

- RF
- log*RRF*
- Signal-conc.

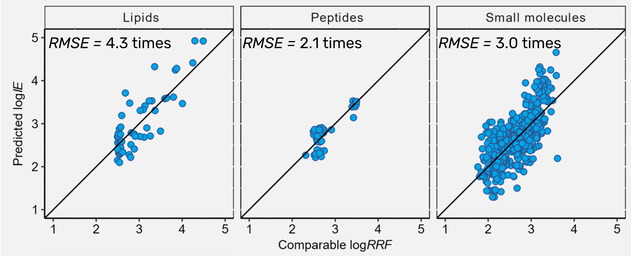
SMILES
Eluent

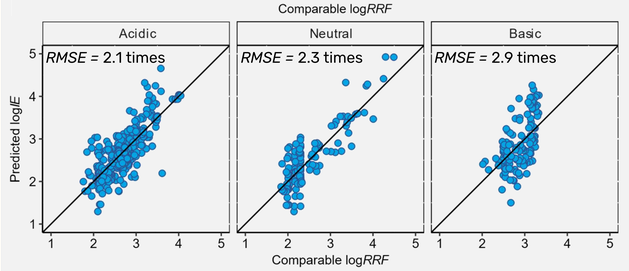
MODEL

$$\log RRF_{comparable} = \frac{\log RRF - intercept}{slope}$$

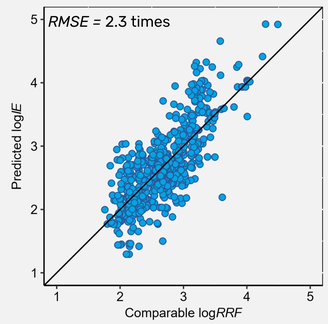


Results for ESI+

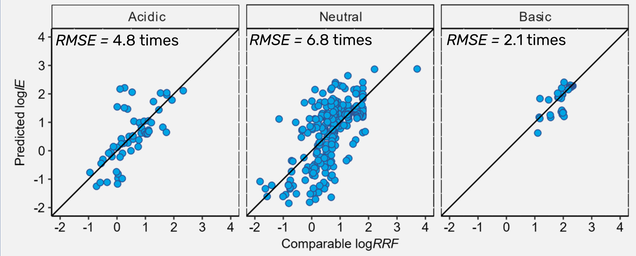




Organic phase:
MeCN and MeOH (RMSE 2 times).

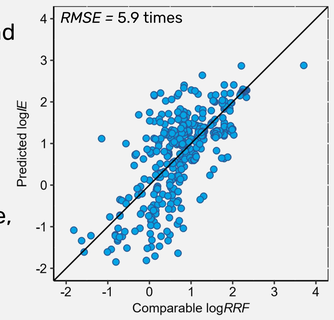


Results for ESI–



No distinctive compound classes in ESI–.

Organic phase:
- MeCN and MeOH,
- Neat solvents (incl. water),
- THF, iPrOH, acetone, chloroform.



Conclusions

- Sharing raw data (calibration graphs, concentrations) can lead to new insight.
- Data in tabulated form is more easily and likely to be used than figures – consider publishing in Supporting Information.
- Approach to **unify literature data** (past, present, future).
- Approach equally good for most used eluent compositions and compound classes.
- Data is consistent - **RMSE 2 times**.
- Leap towards unifying the knowledge about ionization efficiency in ESI/MS community.

K. Stavenhagen, *J. Mass Spectrom.* **2013**, *48*, 627–639.
A. Leitner, *Chromatographia* **2007**, *65*, 649–653.
E. Cifková, *Anal. Chem.* **2012**, *84*, 10064–10070.
W. C. Byrdwell, *J. Chrom. A* **2013**, *1320*, 48–65.
W.-C. Yang, *Anal. Chem.* **2006**, *78*, 4702–4708.
N. B. Cech & C. G. Enke, *Anal. Chem.* **2000**, *72*, 2717–2723.
K. R. Chalcraft, *Anal. Chem.* **2009**, *81*, 2506–2515.
J. Hermans, *Anal. Chem.* **2017**, *89*, 9159–9166.
E. N. Pieke, *Anal. Chim. Acta* **2017**, *975*, 30–41.
C. J. Cramer, *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 278–285.
B. M. Ehrmann, *J. Am. Soc. Mass Spectrom.***2008**, *19*, 719–728.
V. J. Mandra, *Rapid Commun. Mass Spectrom.* **2015**, *29*, 1661–1675.
A. Kiontke, *PLOS ONE* **2016**, *11*, e0167502.
M. Bedner & D. L. Duewer, *Anal. Chem.* **2011**, *83*, 6169–6176.
M. Koivusalo, *J. Lipid. Res.* **2001**, *42*, 663–672.
A. W. Kamga, *J. Am. Soc. Mass Spectrom.***2014**, *25*, 880–890.
C. Monnin, *Rapid Commun. Mass Spectrom.***2018**, *32*, 201–211.
Y. Lin, *J. Chrom. A* **2016**, *1435*, 66–74.
B. A. Huffman, *Anal. Chem.* **2012**, *84*, 9942–9950.
W.-T. Tang, *J. Anal. Methods Chem.* **2014**, *2014*, 1–9.
L. Wu, *Anal. Chim. Acta* **2013**, *794*, 67–75.
T. Henriksen, *J. Am. Soc. Mass Spectrom.***2005**, *16*, 446–455.
C. M. Alymatiri, *Anal. Methods* **2015**, *7*, 10433–10444.
M. A. Raji, *Rapid Commun. Mass Spectrom.* **2009**, *23*, 2221–2232.

PRG300
Eesti Teadusagentuur
Estonian Research Council
European Union
European Regional Development Fund
Investing in your future