**Gradient retention time modeling in ion chromatography through QSRRs and ensemble machine learning**

Petar Žuvela1,\*, Zhen Jia Lim1, Šime Ukić2, Mirjana Novak2, Tomislav Bolanča2, Mario Lovrić3, Ming Wah Wong1, Bogusław Buszewski4

1Department of Chemistry, National University of Singapore, Singapore

2Department of Analytical Chemistry, Faculty of Chemical Engineering and Technology, University of Zagreb, HR-1000 Zagreb, Croatia

3Know-center, Graz, Austria

4Department of Environmental Chemistry and Bioanalytics, Nicolaus Copernicus University, Toruń, Poland

\*Corresponding author. E-mail: [petar.zuvela@nus.edu.sg](mailto:petar.zuvela@nus.edu.sg).

**1. Introduction**

**Notes:**

1) Overcoming the key limitation of QSRR, the validity only for a system of pre-defined conditions.

2) Using ensemble learning to simultaneously model isocratic retention time for all eluent concentrations.

3) Using the developed models to predict gradient retention times using the iso-to-grad model

4) **The iso2grad methodology can be ported to any separation mode, i.e., independent of the polynomial retention factor vs. eluent concentration dependency**.

5) Overcoming the key limitation of QSRR and directly modelling all the isocratic concentrations

6) Notable improvement of predictive ability as omcpared to the previously-developed models (GA-ANNs, GA-ANFIS)

7) Comprehensive comparison of gradient boosting (GB), extreme gradient boosting (xgBoost), adaptive boosting (AdaBoost)

8) Point 4) is particularly important since for sugar alcohols the relationship logk~logc(KOH) is not polynomal but linear.

**Figure 1.** Flowchart of the developed methodology.

**2. Experimental section**

**2.1. IC analyses**

**2.2. Data processing**

**2.3. Model development**

**2.4. Model validation**

**3. Theoretical**

**3.1. Quantitative structure-retention relationships**

**3.2. Ensemble learning**

**3.3.1. Gradient boosting**

**3.3.2. Extreme gradient boosting**

**3.3.3. Adaptive boosting**

**3.3.4. Random forests**

**3.4. Isocratic-to-gradient model**

**3. Results and Discussion**

**4. Conclusions**

**Acknowledgments**

NSCC, ThermoFisher Scientific?