



MultiConditionRT: Predicting liquid chromatography retention time for emerging contaminants for a wide range of eluent compositions and stationary phases

Amina Souihi, Miklos Peter Mohai, Emma Palm, Louise Malm, Anneli Kruve

Department of Environmental and Materials Chemistry, Stockholm University, Svante Arrhenius väg 16, Stockholm 106 91, Sweden



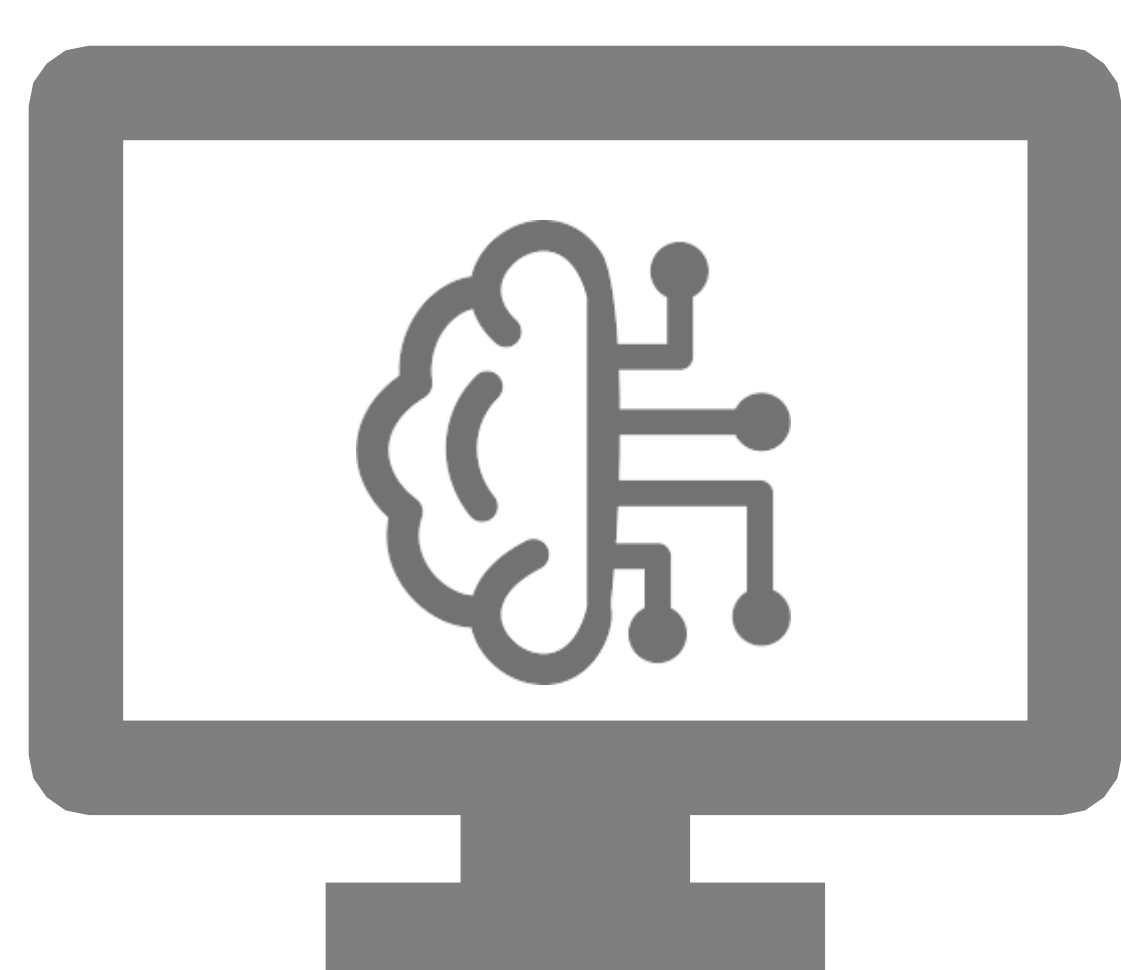
Background

Retention times (t_R) are highly beneficial in the structural elucidation of unknown features in non-targeted LC-ESI-HRMS analysis. In this work, we developed MultiConditionRT for predicting t_R in twenty mobile phases and four columns

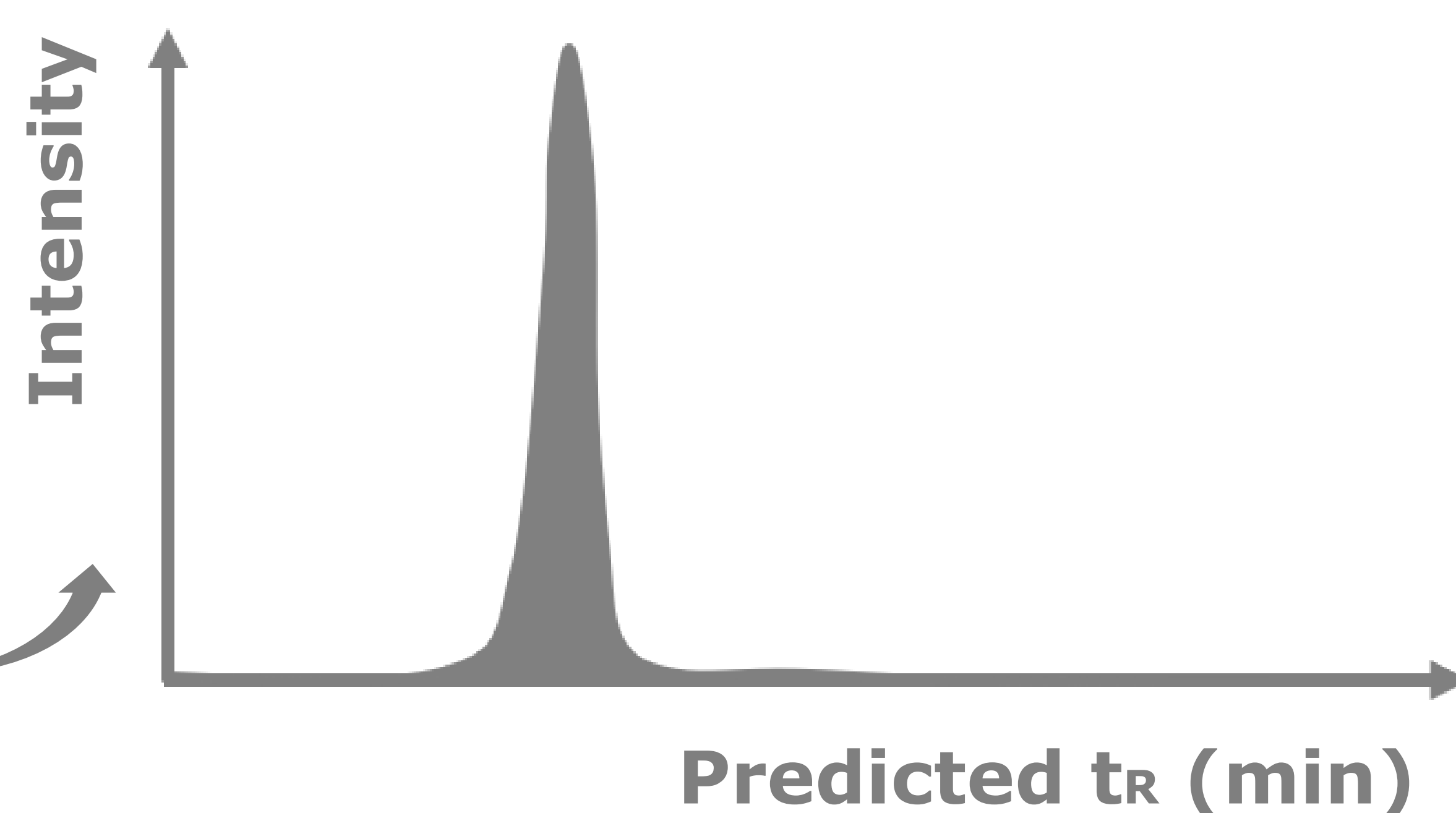


Experimental

Molecular descriptors
 $\log P$, pK_a
Chromatographic conditions



Machine learning



Results

Random forest was the best-performing algorithm

MultiConditionRT predictions could be made in a wide range of chromatographic conditions:



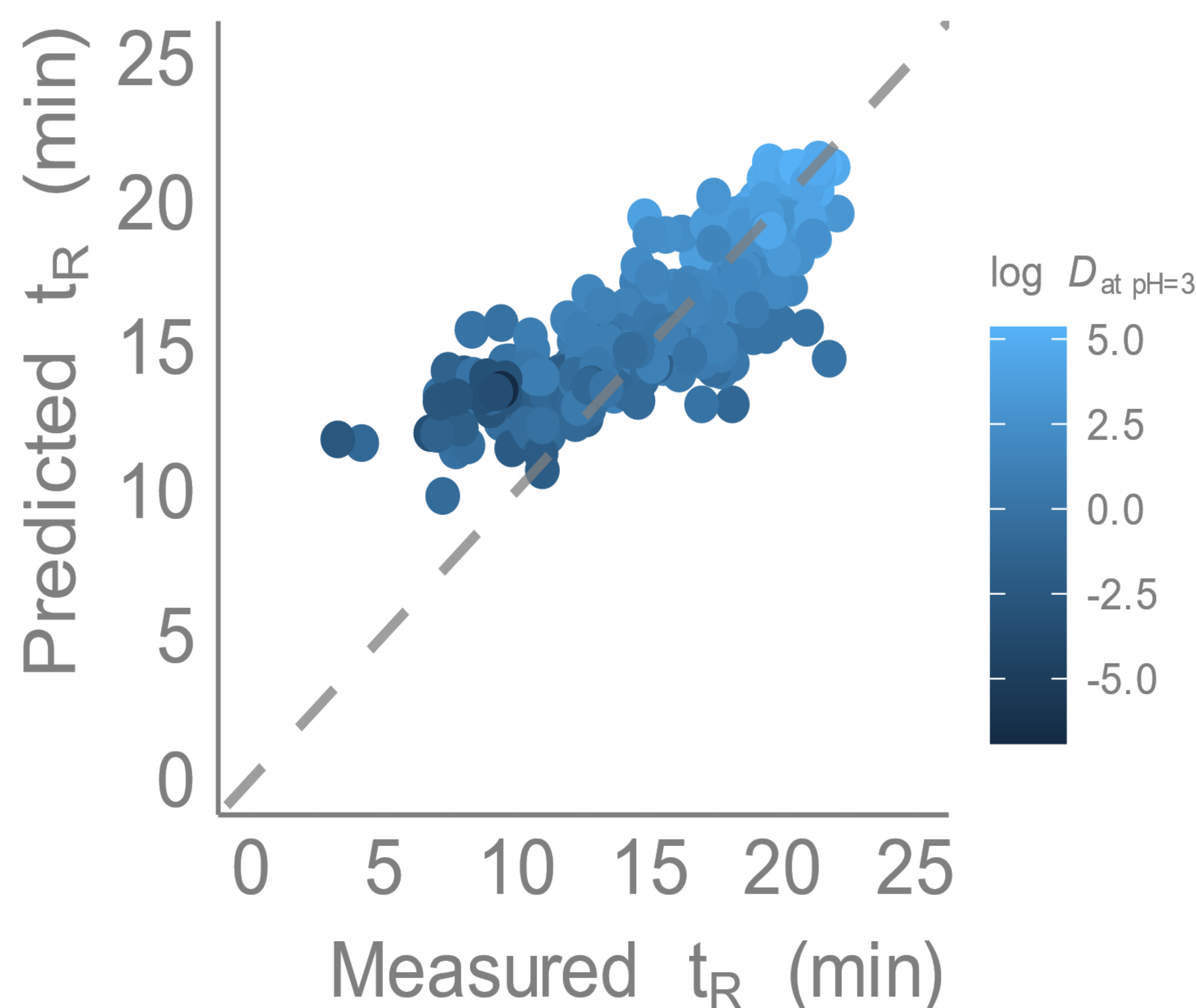
4 Columns: Mixed mode; C18 reversed phase; HILIC and biphenyl



2 organic modifiers: Acetonitrile and methanol



8 water phases: 7 additives and pH ranging from 2.1 to 10.0



The root mean square error of MultiConditionRT is **1.55 min** for C18 reversed phase, **1.79 min** for mixed-mode, **1.93 min** for HILIC, and **1.56 min** for biphenyl column.



Application

In a non-targeted workflow, MultiConditionRT was used to reduce the tentative candidates given by SIRIUS+CSI:FingerID(1). 57.1% of correct structures remained in the candidate list after eliminating 65.2% of tentative candidates

(1) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. SIRIUS 4: A Rapid Tool for Turning Tandem Mass Spectra into Metabolite Structure Information. *Nat Methods* **2019**, 16 (4), 299–302. <https://doi.org/10.1038/s41592-019-0344-8>.