Poster Presentation Preferred

Student Presenter

**Hydrophilic Lipophilic Balanced Retainability in the Context of Chemical Space.** A. Feerick, G. Black, T. Young, University of California – Davis, Davis, CA. Hydrophilic Lipophilic balanced (HLB) cartridges are widely used in solid-phase extraction for sample cleanup and concentration for nontarget analysis. Their ability to retain a diverse set of compounds while simultaneously removing interferences makes them a preferred step in many workflows. Despite the diverse set of compounds that HLB is capable of recovering, there are inevitably many that are not successfully retained. The boundaries of chemical space, i.e., the set of known and possible compounds, covered by HLB extraction remains undefined, limiting the identification confidence of suspect and unknown contaminants annotated during nontarget analytical workflows. Defining the “HLB detectability domain”, an area of chemical space where HLB is capable of extracting versus not based on a multitude of molecular descriptors, is crucial for improving the confidence of feature annotation. We propose the use of computational machine learning-based models to predict the detectability domain of nontarget methods by examining the extraction potential of HLB sorbents. For this purpose, a dataset from a previous study consisting of 414 priority and emerging pollutants in water was used. One- and two-dimensional molecular descriptors for each compound were obtained through PaDeL. Both classification and regression algorithms were evaluated for their potential in developing a quantitative structure-property relationship (QSPR) model that will describe which molecular descriptors most affect HLB retainability. The classification methods include classification and regression trees (CART), random forests (RF), and genetic algorithm (GA)-support vector machines. Both linear (GA-multiple linear regression) and non-linear (GA-support vector regression) regression methods were employed. RF classification models were the best performing based on their kappa value. Regression models were not successful in predicting a chemicals percent recovery, likely because the data set included a limited number of values within some recovery ranges. Knowing the range of compounds that can be readily recovered by HLB resins, confines nontarget annotations to this area of chemical space. Understanding the bounds of HLB extraction (nontarget space) is an important step in nontarget method standardization and reproducibility, both of which are necessary for improving regulatory acceptance of the methods for environmental monitoring and assessment.