Analysis of Two-Dimensional Gas Chromatography Data

Presented by DSTL (Alex Harvey and Emily Matthews)

Gas chromatography (GC) is a popular tool for chemical analysis. In GC, a sample is passed through a column and the analytes within the sample are retained by the column depending on their affinity to the column phase. Therefore each analyte will move through the column at different rates and reach the detector where a retention time is recorded. Samples such as air, soil and other environmental samples are so complex that a single column does not have enough power to separate all of the analytes. In this instance a higher resolution GC method, known as comprehensive 2-dimensional gas chromatography (GCxGC), is used. GC x GC combines two columns, hence two retention times and the volume of each analyte is used for identification. GCxGC allows for less than pico-gram detection of the chemicals in a sample.

We want to be able to use data from GCxGC to identify features of the background of a sample, with the aim of attributing samples to a particular region or cultivar. In particular, we have GCxGC data for multiple seeds, which are of the same type but come from different countries and cultivars. Our aim is to use the GCxGC data to identify similarities between seeds from the same country and cultivar, if they exist.

Collecting the data from these different seeds for comparison, and identifying appropriate methods for this comparison, is a complex problem. There is a suggestion that the Bayesian methods of Gabriel Vivo-Truyols and co-authors is the solution, however alternative solutions may also allow us to analyse this data without losing information.

Dstl are seeking methods or tools that both collects the data from the different samples and analyses this collected data to identify similarities between samples from the same countries or cultivar. This solution should be able to be run on a standard laptop, and preferably should be written in C++, R, Matlab or Python.

References

Vivo-Truyols, G. (2012). “Bayesian Approach for Peak Detection in Two-Dimensional Chromatography.” Anal. Chem., 84 (6), pp 2622-2630.

Barcaru, A. and Vivo-Truyols, G. (2016). “Use of Bayesian Statistics for Pairwise Comparison of Megavariate Data Sets: Extracting Meaningful Differences between GCxGC-MS Chromatograms Using Jensen-Shannon Divergence.” Anal. Chem., 88 (4), pp. 2096-2104.

Barcaru, A., Derks, E. and Vivo-Truyols, G. (2016). “Bayesian peak tracking: A novel probabilistic approach to match GC x GC chromatograms.” Analytica Chimica Acta, 940, pp 46-55.