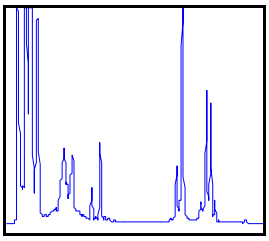

InfoMetrix®

Product Note

Automating the Interpretation of Chromatograms

Chromatography is an extremely versatile technique for the analytical or quality control laboratory. As such, chromatograms are collected for a wide variety of both quantitative and qualitative assessments. In many cases, the interpretation of chromatographic output is subject to manual interpretation. We rely on holding an overlay of two chromatograms up to a window to see if a process or product has undergone "significant" change. Chemometrics affords us the ability to substitute an objective and automated mechanism for the human pattern recognition step. We do this to provide reliable answers immediately upon completion of the instrument run and further afford the skilled analyst the chance to ply his or her trade on the problems for which the automated engine has insufficient information.

The Infometrix software Pirouette® and InStep™ allow users to make chromatographic pattern matching systems without much customization. These products build on the experience gained from building several custom systems, spanning HPLC, GC and GC/MS for the FDA, the CDC and several food and petroleum companies, automating interpretation for instruments from Agilent, PE, Waters, Shimadzu and others working with GC, HPLC, HPCE, TLC and gel electrophoresis.

The steps:

1. Collect the experience set;
2. Build the Pirouette models;
3. Set up the InStep method and reports; and
4. Configure the instrument software to save files into a watched folder.

Use of Peak Table or Whole Chromatogram

In essence, there are two ways to look at the output of a chromatograph: use the peak table generated by the chromatographic software or look at the chromatogram as if it were a spectrum (*i.e.*, use the whole chromatographic profile).

Peak tables have to be constrained in order to work in the automated world. The peaks making up the signature have to be the same peaks from analysis to analysis and suffer from two problems: retention time shifting leading to misidentification and the fact that any unexpected peak would not be processed. From our experience, the primary problem is the unexpected peak.

Using the whole profile solves the unexpected peak problem and handles shoulders and other chromatographic features, but it, like in spectroscopy, can suffer significantly from x-axis (in this case retention time) instability. Pirouette has built-in the ability to correct for retention drift, dramatically improving the quality of whole-profile matching. Files with the retention times corrected can be saved.

Applications

Batch Quality Control

The manufacture of synthetic biologicals is treated as a batch process. Because of the complexity of the product, quality control is conducted using capillary electrophoresis. The pattern is complex and the majority of the fragment peaks are not identified. There are changes in the chromatogram or electropherogram from batch to batch. Pattern recognition technology provides a simple means to determine what variation is acceptable (no impact on

the efficacy or the safety) in the case where the tolerance for variability is not uniform for all peaks. In other applications:

- Residual solvent patterns can detect process changes
- Product quality can be graded based on more than just the "favorite ratio"
- QC runs can be corrected for retention time shift prior to archival

Product Origin

A state Department of Justice (DOJ) has built a database for comparison of complex chromatograms obtained from clandestine lab (methamphetamine) samples. Their knowledge base consists of GC or GC/MSD data from 100 laboratories in the state. When a new sample is found, it is run against the database to see if there is a match. From the chromatographic pattern, they gain information about the source and the process used to manufacture the illegal drug. Other applications have used:

- Process GCMS to characterize PCB distribution and search against Aroclor patterns
- Use GC/FID to identify the source of spills in a refinery
- HPLC fingerprinting of herbs, fruits, etc. to determine origin or shipping history

Mining an Experience Set

A food company has a continual need to check samples for identity and compare samples with standards for quality control or supports of other applications. From a library of 3,000 chromatograms, Pirouette was used to establish an optimal, non-redundant database of good QC samples. Now in routine use, InStep evaluates every new sample; when the software flags the product as being out of class, then Pirouette is used to see why. Additional work has shown that:

- *M. tuberculosis* can be identified from a database containing 60 related species
- Source rocks for any crude oil can be typed based on GC and GC/MS
- Chromatograms collected over a multi-year timeframe can help classify aberrant samples

Quantitating Change

A major application for this software is the comparison of peptide maps resulting from the enzymatic digestion of recombinant proteins. Peptide mapping is a sensitive technique that biotechnology companies use to assess the quality of their protein products. Chromatographic pattern recognition allows detection of 1% or smaller changes in the composition of the products, thereby helping to meet guidelines for well-characterized biopharmaceuticals. Similar applications have demonstrated:

- Quantitation of the degree of weathering for fuel classes in arson investigation
- GC column QC to identify problems before they cause data loss
- Identification of seasonal variation in blended products
- Differentiation between a chromatographic problem (all of the significant variables are expected to shift) and an impurity problem (only some of the variable shift)

Mixture Analysis

A spice that is a mixture of various natural ingredients can pose the problem that the product is supposed to stay consistent from year-to-year, but the ingredients can vary. What is the best mixture to reproduce the standard product? Can a new ingredient (source or material) be substituted in the formulation? Mixture analysis also applies to:

- Quantitating relative composition after backblending kerosene into diesel #2
- Identifying end members of mixtures (binary, ternary +) based on the GC pattern
- GCMS peaks can be unmixed to quantitate impurities, even in the absence of standards
- Deconvolute HPLC peaks based on UV spectra, GC peaks based on MS spectra

How to Start

Obtain Pirouette and InStep by purchasing a CD or downloading from the Infometrix web site. Develop your own methods using the algorithms in the software or work with Infometrix to create a custom solution for your application.

Infometrix, Inc., 10634 E. Riverside Dr., Bothell, WA 98011 Phone: (425) 402-1450, Fax: (425) 402-1040, email: info@infometrix.com