Pacific Northwest NATIONAL LABORATORY

*Proudly Operated by* **Battelle** *Since 1965*

Battelle Number(s):

12120-E, 12836-E, 13902-E

Available for licensing in all fields

U.S. DEPARTMENT OF ENERGY

Available Technologies

# AlgoWorks - New Approaches for Processing Spectral Data

## SUMMARY


Fully Automated Analysis

An important problem in analysis of spectral data is identifying and characterizing significant features. Traditionally, peak detection is based upon finding signals above an arbitrary threshold value, such that peak identification is affected by variable signal to noise ratio, and varying baseline signal. The results are difficult to reproduce, suffer from substantial signal loss, and are subject to uncertainty.

The method developed at PNNL for peak extraction is fully automated and defines a method of processing data that is fundamentally different from other methods. Whereas other methods focus on looking at the spectral peak height (the y-value), this new method uses a combination of x and y values to analyze the spectral data. The result is a method of analyzing spectral data that

* Is significantly less sensitive to signal to noise ratio, signal drift and varying baseline signal;

* Allows for high throughput, automated data analysis;

* Provides greater confidence in correct identification of spectral peaks;

* Is successfully being applied to the analysis of mass spectral, gas chromatography and nuclear magnetic resonance spectral data showing its broad applicability

Process Control of Mass Spectral Data

An increase in applications of high throughput and/or automated mass spectral analysis has generated an increasing need for quality control algorithms that can reliably and accurately indicate instrument degradation or failure, or possible sample contamination. The traditional chemometrics approach to process control of spectral data relies solely on peak intensities, and may be inappropriate for some applications in mass spectrometry, where peak presence is more important than peak intensity.

PNNL researchers have developed a statistical model through which mass spectral data can be analyzed. This model describes the statistical properties of a spectrum using the rules of conditioning, and incorporates discrete and continuous multivariate statistical methods. This model is general enough to handle mass spectral data, where peak locations are of primary interest but intensities are not. It can also handle other instrument data, where both peak locations and relative intensities may be of interest. From this new model, a cumulative sum (CUSUM) test for quality control of spectral data is derived. Based on a likelihood ratio hypothesis test, the CUSUM test sequentially compares test samples to a reference fingerprint comprised of normal control spectra. The test samples are considered normal until the CUSUM test statistic exceeds a critical threshold, at which time the samples are considered abnormal.

A New Scoring Method for Peptide Identification via Tandem Mass Spectrometry

PNNL Researchers have also developed a new approach for peptide identification from a list of candidates using collision-induced dissociation tandem mass spectrometry data. This approach is based on a probabilistic model for the occurrence of spectral peaks corresponding to frequently observed partial peptide fragment ions. As part of the identification procedure, a probability score is produced that indicates the likelihood of any given candidate being the correct match. The performance of the proposed algorithm has been tested on a large test dataset consisting of spectra from tryptic peptides and with different lengths ranging from 6-mers to 30-mers. Results suggest that this approach works well with error rates typically below 5 percent.

## Technology Portfolio(s)

» Biotechnology

## Potential Industry Applications

» Chemicals

**Bruce J.** Harrer
Pacific Northwest National Laboratory
(509) 375-6958
bruce.harrer@pnnl.gov
http://availabletechnologies.pnnl.gov

Pacific Northwest
NATIONAL LABORATORY

*Proudly Operated by* Battelle *Since 1965*