KnowItAll® AnalyzeIt MVP

Multivariate Processing Made Simple

AnalyzeIt™ MVP, which incorporates Infometrix' chemometrics technology for principal component analysis (from the well-known Pirouette® software), is a complementary addition to Bio-Rad's KnowItAll® data analysis environment.

The integration provides a powerful tool for expert and nonexpert users alike to perform multivariate analysis of spectroscopic, chromatographic, or numeric data.

Benefits

- Gain insight into hidden patterns / relationships in data
- Explore data correlations to answer critical research, development, or production questions
- Facilitate the storage of analysis results for subsequent reference, reporting, or investigation

What is Multivariate Analysis?

Multivariate analysis (including principal component analysis, PCA) refers to the statistical analysis techniques where multiple variables are analyzed to determine the contribution made by each variable to an observed result. This permits patterns to emerge from within the data. Researchers can use this method of analysis to examine quantitative data in more depth than from a basic cross-analysis of the data.

AnalyzeIt MVP: IR Example

PCA was performed on an IR database of automobile paint chips. The legend at the right is also an integrated Query Device, allowing the user to select one or more paint classes by clicking on the appropriate row or rows in the legend—clicking on the "Water Based Enamel" row in the "Paint Type" legend highlights all points in the scatter plot and displays all spectra that correspond to the paint type selected in the legend.
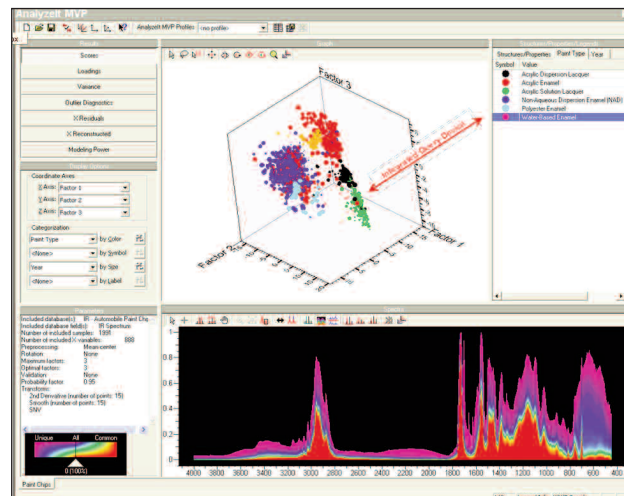
AnalyzeIt MVP: NMR Example

PCA was performed on an NMR database of organic compounds with a diverse range of functional groups. Note that structures can be made visible via the "Structures/Properties" pane, which allows immediate visualization of the structures associated with selected data points. Also, the Overlap Density Heatmap slider bar has been moved toward the "common" side to show only those spectral areas with higher levels of similarity.

BIO-RAD
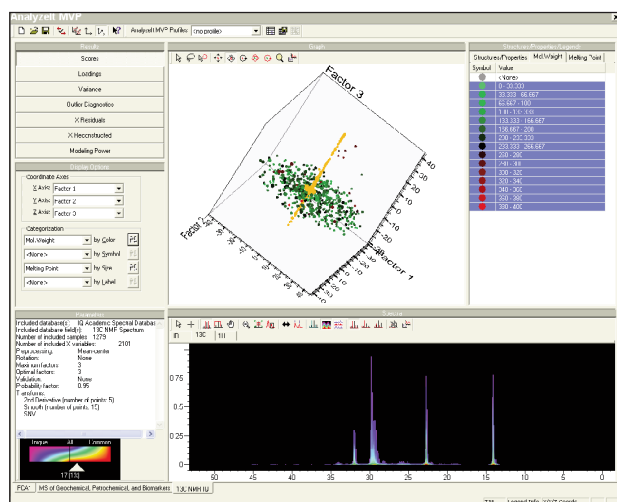
Features

Integrated Data Import, Management, Pre-Processing

As part of the KnowItAll Informatics System, AnalyzeIt MVP takes advantage of all of the additional applications in the KnowItAll environment, including data import and processing of spectral and chromatographic data via batch mode; database creation, visualization, searching; and report generation. In addition, the complete list of data preprocessing options familiar to Infometrix' Pirouette customers is incorporated into AnalyzeIt MVP, allowing complete flexibility in the multivariate analyses performed. Pre-processing options for PCA include:

- **Pre-Processing** (Autoscale, Mean-Center, Pareto, Range Scale, or Variance Scale)
- **Y-Transforms** (1st Derivative, 2nd Derivative, log10, Multiply, Subtract, MSC, SNV, Smooth, Normalize, and Baseline Correction)
- **Binning/Bucketing** (Fixed Width, IntelliBucket™ or User-Defined Range)
- Project a standard database to the PCA space of a collection of data

Plots, Spectra, Structures, and Legends as Integrated Query Devices

The plot legends in AnalyzeIt MVP double as simple but powerful query devices, turning the application into a data mining tool. The legends, plots, spectra, and structures are fully interactive and interlinked:

- If the user selects points in the scatter plot, all corresponding spectra and structures be displayed and all rows in all legends corresponding to the selected points will be highlighted.

- If the user selects a row in any legend, all points corresponding to that legend item will be highlighted in the graph, all corresponding spectra and structures in the data set will be displayed, and most importantly, all corresponding rows in other legends will also be highlighted.

Spectral displays can then be viewed with Bio-Rad's patented Overlap Density Heatmap (ODH) technology for comparative visualization.

Display Options

AnalyzeIt MVP also allows one to draw the time-trajectory course of data over time. Several options for datapoint display are available, with intuitive default values that still give the user complete control to customize all plot display features, including categorization by:

- **Color** - Change colors in continuous or categorical coloring schemes
- **Symbol** - Plot different categories of data using different symbols
- **Size** - Plot different categories or ranges of data with different size symbols
- **Label** - Display a data label, including selectively labeled points according to user-specified data ranges

Post-PCA Interpretation Tools

- **Use loadings plot as query to search database:** The resulting loadings plots and the OD consensus spectra of different classes or their difference spectra can be used as a query spectrum to search against spectral databases in KnowItAll.

- **Project reference database on samples' PCA:** This method complements the above approach in that it projects a "known" reference database of spectra or chromatograms to the PCA space of a sample collection and therefore, uses the known record's position to identify the unknown "marker" of the sample. (See datasheet on Database Projection Analysis for more details).