MS2 Data Mining & Visualisation
Project Update

AJM Consulting

Issue 4
September 2007

Real-time Diagnostics Project Starts

Now that the benefits of analysing existing data for completed production have been proven on many client sites, we are responding to client requests by extending our development to include real time analysis of current production.

As was the case with the existing system's development, we are using a partner-based approach which means that the design is highly client-driven. Two companies are already on board with the design phase; ConocoPhillips and Syngenta. There is still a limited opportunity for others to join, which will result in these companies getting the system at a much reduced cost yet with their specific requirements addressed by the design.

The new project (codenamed Voyager) will result in a system which will automatically acquire data from plant systems such as a DCS or process historian, and compare this data to models of known good performance built using the existing MS2 system.

As with the existing MS2, ease of use and flexibility will be key aspects of the design which will be applicable to both batch and continuous processes. In addition to modified versions of existing algorithms such as Principal Component Analysis, others will be developed for aspects including acquisition and storage of data, analysis and visualisation.

The goal of Project Voyager is the realisation of real benefits including:

- Early indications of subtle shifts in process conditions or of faults developing
- Warning of movement into operating parameters which eventually cause plant damage
- Early predictions of loss of quality or yield

South West office facilitates expansion into bioscience

MS2 can be applied to many data sets, not just those relating to chemical manufacture. For instance it is being used in a bioscience industry cluster mapping study which we are providing to South West Regional Development Agency (SWRDA). The south west has a major cluster of bioscience industries and, to facilitate MS2's growth in this area, AJM Consulting has opened a new office in the region, located at Plymouth's prestigious Tamar Science Park.

Tamar Science Park

Inside this issue:

Development of the MS2 Process Analysis System received funding assistance from DTI's Research & Development Fund, administered by Yorkshire Forward.

We are particularly pleased to be working closely with the Centre for Process Analytics and Control Technology at Newcastle University and are grateful for the assistance we receive from this renowned centre of multivariate research and development.

Progress Continues

With 30 users, new industry sectors, a new office and now overseas marketing, we face interesting times!

Welcome to the fourth issue of our data mining project newsletter. There are now over 30 users of MS2, in a diverse range of industries including chemicals, oil, pharmaceuticals, nuclear and, totally differently, public sector, with a bioscience cluster mapping study which the South West RDA has commissioned AJM Consulting to provide, using MS2 to analyse complex industry data.

Our new office in Plymouth becomes operational in October, which will focus particularly on the bioscience industries.

With the imminent release of the latest version, MS2's functionality continues to expand to meet the needs of the diverse applications to which it is now applied.

Training courses have proven very popular. With all four courses fully booked prior to the summer, those scheduled for October and November are already heavily booked.

Functionality continues to grow; for instance we have integrated Manhattan significant change detection into the system; this arose from a competition-winning suggestion made at the first training course. It was well worth the champagne!

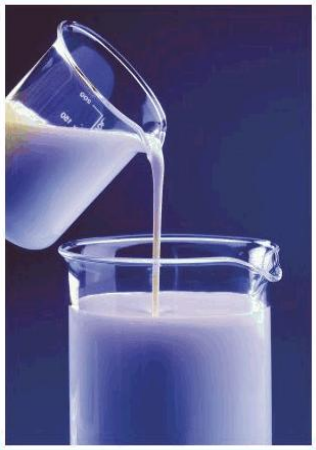
Also during the summer we exhibited overseas for the first time, at ChemSpec in Amsterdam, and as a consequence are in discussions with several overseas companies in Europe and North America.

The real time diagnostics project has now started its design phase, with involvement from two companies (ConocoPhillips and Syngenta).

Alan Mason
Managing Director

Case Study - Synthomer

Synthomer is a world-class supplier of synthetic polymers to industries ranging from paints and adhesives to textiles, speciality papers and plastics. At its manufacturing plant at Stallingborough, it produces a wide range of synthetic latex products which are used in many types of industrial applications. The company has grown rapidly in recent years by being faster and more innovative than its competitors and quality is paramount.



Achieving this consistently high quality means understanding the variability inherent in a modern, complex manufacturing process, and Synthomer selected the MS2 Process Analytics System from AJM Consulting to assist in the identification of causes of variability between multiple reaction vessels.

The various computer systems on site hold extensive sets of data relating to quality parameters, batch operations and process trends within reactors. To fully visualise the manufacturing characteristics of each batch it was necessary to integrate these data sets and the MS2

"We initially asked AJM Consulting to investigate a process issue in which two reactors had differing characteristics which we did not fully understand. We had plenty of data, but interpreting it was a problem because it was derived from several systems and could not be integrated. The MS2 system integrated the data from our commercial, process control and plant historian systems and provided both an effective visualisation of the problem and an analysis of its cause, enabling us to implement changes which have resulted in increased consistency and quality."

"We are now implementing MS2 across the plant."

Stuart Askham
Synthomer Ltd

Version 4.4 provides many new features

The latest version of the MS2 data mining system will be released shortly. This version contains many new features which have been developed to meet the requirements of the 30 systems now in use at client sites. There are also a host of other improvements to correct known faults and increase both visualisation and ease of use.

Significant Change Analysis

Prominent amongst the new features is significant change analysis using the Manhattan algorithm. Whilst significant change points in a time series plot can be analysed to an extent by the CUSUM function, this approach is imprecise. With Manhattan, each change is automatically identified. Within MS2, Manhattan can be applied to any variable, such as a primary process variable or the score of a principal component. Manhattan is also totally integrated with the other features of the MS2 system.

Batch Trends

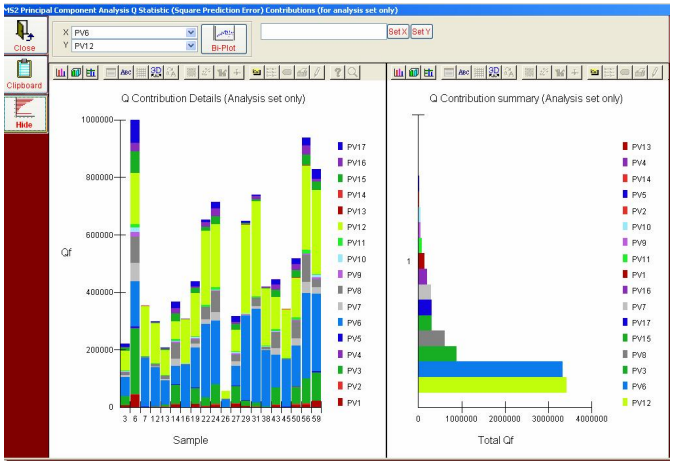
Formerly available as an optional extra, the MS2 batch trend pre-screening functions have been both upgraded and made available as a standard part of the MS2 system. Trend curves (for example temperature and pressure) are very important data sets for variability diagnosis, but analysing them can be a difficult task. MS2 now enables each batch's trend files to be checked for errors which would preclude multivariate analysis, converted to a standard file size (a requirement of principal component analysis) and imported into MS2.

Square Prediction Error Analysis

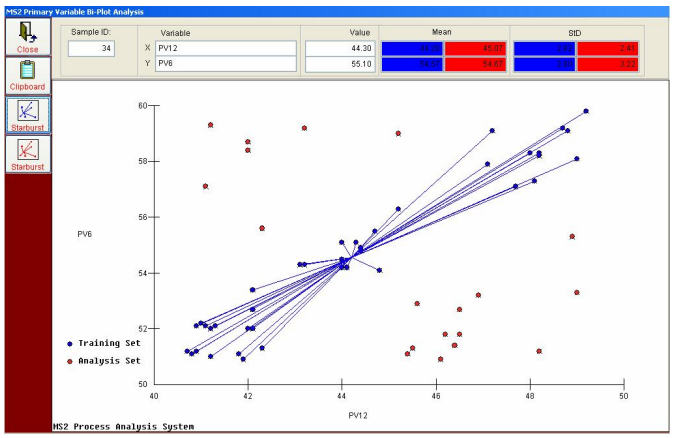
The Q statistic (square prediction error) has always been calculated by MS2's multivariate analysis functionality, which now includes drilldown to visualise the contributions to the value of Q. This provides a very powerful way of identifying subtle multivariate links between variables.

Screening Rules

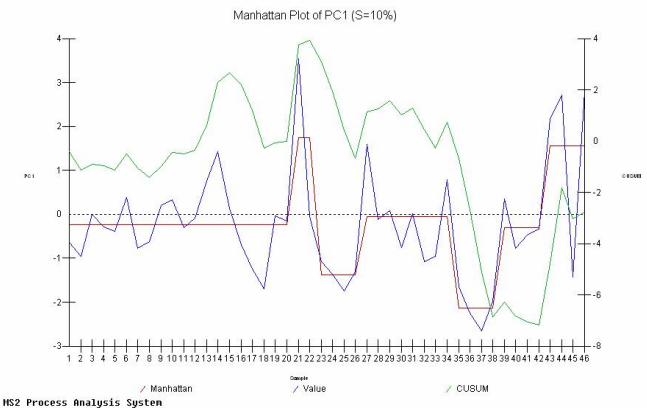
The screening functionality in MS2 has always been appreciated for its ability to prepare raw data sets for analysis easily and, where those sets change with time, repetitively. Now the rule set has been enhanced with some new rules, for instance several types of noise reduction. Improvements have also been made to the CLASSIFY rule which enables non-numeric data to be analysed.



This data set with 17 variables shows that two have much greater contributions to the Q statistic than the remainder.



Orthogonally plotting the two variables identified in the Q Statistic contribution plot above shows a clear multivariate relationship (blue) which, if not followed, results in poor quality (red).



An example of Manhattan (red) applied to a PC score (blue). The CUSUM plot (green) is far less distinct at identifying significant change.

AJM Consulting

Europarc Innovation Centre
Europarc
Grimsby
North East Lincolnshire
DN37 9TT

Phone: +44 (0)1472 500306
Fax: +44 (0)1472 500307
E-mail: cathy.oswin@ajm.co.uk
http://www.ajm.co.uk

Creating Manufacturing Advantage

AJM Consulting is a member of:

- *Institute of Measurement and Control*
- *Humber Chemical Focus*
- *Yorkshire Chemical Focus*
- *NEPIC*
- *CPACT*

Our Clients include:

- ABB
- Arthur D Little
- AEA Technology
- Aesica Pharmaceuticals
- A H Marks
- BP
- British Sugar
- British Paper & Board
- Cambridge Consultants
- Carbon Trust
- ConocoPhillips
- DEFRA
- DICIDA
- DTI
- Eli Lilly
- Fibres Worldwide
- GlaxoSmithKline
- Glinojeck Sugar Refinery
- Grotech Production
- Hickson
- Holliday Pigments
- Humber Chemical Focus
- Huntsman
- ICI
- Impress/HCCTA
- Imerys
- J R Crompton
- Kemfine
- Lakeland Laboratories
- Luxus
- Manro
- Norsk Hydro
- PICME
- PSE
- Roche
- Rotork
- Scott Bader
- Servelec
- Severn Trent Water
- SIRA
- Spiritus Consulting
- Strategem Consulting
- Stepan
- STG
- Sulzer
- SWRDA
- Syngenta
- Synthomer
- Technical Absorbents
- Tensachem
- Total
- University of Newcastle
- University of Sheffield
- University of Strathclyde
- UMIST
- Yorkshire Forward

...and others we can't mention

Frequently Asked Questions

What are the forward plans for development after the current release?

We will release Version 4.4 during September 2007 and the next version after that will be V4.5, which will be released in early 2008.

Do these releases contain new functions or just improvements to existing ones?

They contain new functions. For instance, V4.4 includes Manhattan significant change analysis which is highly integrated into the system. V4.5 also has new functions, which are currently in design, but you'll have to wait and see what they are!

Will MS2 work with Microsoft's new Vista operating system?

Of course! Now that Vista is fully released we have, of course, tested both installation and operation and it works very well.

Which Windows versions does MS2 require?

MS2 is compatible with Windows 2000, XP(SP2) and Vista. For some functions, the .NET 2,0 Framework and MS Word must be installed.

What are the different functionality levels of MS2?

There are now three levels of functionality provided as standard within MS2:

- Express Edition (formerly known as Level 2) provides data import and screening functions, together with uni-variate data analysis using parallel coordinate visualization.
- Professional Edition (formerly known as Level 3) also includes multi-variate functions including Principal Component Analysis
- Premier Edition (formerly known as Level 4) also includes Multiway PCA batch trend analysis and Manhattan significant change detection.

Now that many companies use MS2, what business benefits have they gained?

Profit through greater understanding of variability. Problems to which MS2 has been applied include batch to batch variability, conditions which cause problems such as poisoned catalysts or fouled heat exchangers, shift performance, raw material characteristics and many others.

Can MS2 be applied to my data?

Almost certainly yes. Originally designed for chemical applications, MS2 is now also used in many manufacturing sectors including oil, pharmaceuticals, surface coatings, advanced materials and nuclear. It is also used to analyse the contribution of the bioscience cluster to the economy of the South West Region! We also envisage many more applications such as clinical trial analysis.